Efficient Randomised Broadcasting in Random Regular Networks with Applications in Peer-to-Peer Systems *

Petra Berenbrink
School of Computing Science
Simon Fraser University
Burnaby B.C. V5A 1S6, Canada
petra@cs.sfu.ca

Robert Elsässer
Department of Computing Science
University of Salzburg
5020 Salzburg, Austria
elsa@cosy.sbg.ac.at

Tom Friedetzky
School of Engineering and Computing Science
Durham University
Durham, DH1 3LE, U.K.
tom.friedetzky@dur.ac.uk

Abstract

We consider broadcasting in random d-regular graphs by using a simple modification of the *random phone call model* introduced by Karp et al. [25]. In the phone call model, in every time step, each node calls a randomly chosen neighbour to establish a communication channel to this node. The communication channels can then be used bi-directionally to transmit messages. We show that, if we allow every node to choose *four distinct neighbours* instead of one, then the average number of message transmissions per node required to broadcast a message efficiently decreases exponentially. Formally, we present an algorithm that has time complexity $O(\log n)$ and uses $O(n \log \log n)$ transmissions per message. In contrast, we show for the standard model that every distributed algorithm in a restricted address-oblivious model that broadcasts a message in time $O(\log n)$ requires $\Omega(n \log n / \log d)$ message transmissions.

Our algorithm efficiently handles limited communication failures, only requires rough estimates of the number of nodes, and is robust against limited changes in the size of the network. Our results have applications in peer-to-peer networks and replicated databases.

1 Introduction

We consider the problem of dynamic broadcasting in random networks with small degree. Broadcasting is one of the most useful, versatile and well-studied communication primitives in distributed computing with many applications, e.g., the maintenance of replicated databases [7], where updates made at some of the nodes need to be propagated to all the nodes in the network. To ensure that all copies of the database converge to the same content, efficient broadcasting algorithms are crucial. Our interest in random regular networks is motivated by peer-to-peer (P2P) systems. The idea of using random graphs to build overlays for P2P systems appears in e.g. the Gnutella network [21] and JXTA of Sun Microsystems [2]. Important properties of P2P networks include connectivity, low degree, high expansion and small diameter. These properties are perfectly fulfilled by the random

*Preliminary version published in the Proceedings of the Twenty-Seventh Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC 2008).

regular graphs considered in this paper. Random topologies with small degree naturally arise in P2P systems, in which overlays are generated according to a Markov process intended to construct and maintain a network with the properties described above (e.g. [32]). Several research groups have recently designed a variety of "random-like" networks for P2P systems (e.g. WARP of [24]), and there is a considerable amount of work devoted to the generation and maintenance of random regular graphs (e.g. [5, 16, 27, 29]).

Since P2P networks are important decentralised platforms for sharing data and computing resources, it is extremely important to provide efficient, simple and robust broadcasting algorithms for P2P overlays. One of the simplest communication models for broadcasting is the so-called *random phone call model* introduced in [25]. In this model a graph is given. In any discrete time step the nodes of the graph can create messages each of which needs to be transmitted to all nodes of the graph. Then (again in every step) each node v calls a randomly chosen neighbour w to establish a communication channel to w. The channel can be used for bi-directional communication during that step, meaning that v can send messages to w and vice versa. The nodes can combine messages, and they can also transmit via several open channels in one step. At the end of the step all open channels are closed. The major drawback of the phone call model is that a node establishes communication channels without knowing if there is any (as yet unknown) message in the system, and which of the messages received so far by this node are already known to its neighbours. This means that many unnecessary communication channels will be established between the nodes if only very few messages are created in a step. This model is therefore of particular interest in situations where messages are generated with high frequency. Then the cost of establishing communication amortises nicely over all transmissions. This, in turn, means that for the analysis it is possible to consider the number of transmissions for each single message separately (see [25]).

In the phone call model we distinguish between *push* and *pull* transmissions, depending on the direction in which the message is forwarded (cf. [25]). In the case of push transmissions, calling nodes send their messages to their neighbours, whereas in the pull model messages are transmitted from the called node to the calling one [7]. Karp et al. [25] noted that, in complete networks, the pull model is inferior to the push model until roughly $n/2$ nodes are informed (i.e. received the message), and then the pull model becomes more effective. As we will see in the next paragraph, the combined push and pull model of [25] is able to broadcast in time $\log_3 n + O(\log\log n)$ with $O(n\log\log n)$ message transmissions, whereas the push algorithm from e.g. [7, 33] uses $\Theta(\log n)$ time and $\Theta(n\log n)$ message transmissions.

To compare the push and pull approaches in more detail, let us for a moment consider the distribution of a single message in a complete graph. In the push model, the number of informed nodes grows exponentially in the first phase, that is, as long as fewer than $n/2$ nodes are informed. From then on, in the second phase, the number of non-informed nodes decreases by a constant factor during every round. In the first phase, $O(n)$ messages are transmitted. The second phase requires time $\Theta(\log n)$ and thus $\Theta(n\log n)$ message transmissions, w.h.p.[1] Hence, the push model requires $\Theta(\log n)$ steps and $\Theta(n\log n)$ message transmissions in complete graphs (cf. [25]). In the pull model, in the second phase messages are spread faster. As soon as $n/2$ nodes of a complete graph are informed, every node becomes informed within $O(\log\log n)$ additional rounds ([25]) and thus only $O(n\log\log n)$ messages are needed. The drawback of the pull model is that in the first phase the node creating a message may have to wait for some number of rounds until it is called for the first time. However, the first phase takes time $O(\log n)$ and uses $O(n)$ messages until $n/2$ nodes are informed, which is asymptotically still the same as in the push model. This implies that a total of at most $O(n\log\log n)$ transmissions is sufficient if broadcasting is stopped at the right

[1] By *w.h.p.* or *with high probability* we mean with a probability of at least $1 - 1/n^{\Omega(1)}$.

time.

In this paper we consider broadcasting in d-regular graphs. We are especially interested in graphs with small degrees, and we consider the number of steps and the number of transmissions induced by a broadcast message. Our goal is to develop time-efficient broadcasting algorithms which can handle (limited) communication failures in the network, as well as changes of the network's size and topology, and which produce a minimal number of message transmissions. The first point is important as the structure of P2P networks changes dynamically due to clients joining or leaving the network. Message minimisation is important for applications such as the maintenance of replicated databases where often huge amounts of broadcasts are necessary to deal with frequent updates in the system.

1.1 Related Work

Most papers dealing with randomised broadcasting analyse the run-time of the push algorithm in different graph classes. Frieze and Grimmett show that, with probability $1 - o(1)$, it is possible to broadcast a message in a complete graph on n nodes in time $\log_2(n) + \ln(n) + o(\log n)$ [19]. Later, Pittel improves this bound to $\log_2(n) + \ln(n) + O(1)$ [33]. In [17], Feige et al. determine asymptotically optimal upper bounds for the run-time of the push algorithm in $G_{n,p}$ graphs (the traditional Erdös-Rényi random graphs [14, 15]), bounded degree graphs, and Hypercubes. In [20] Fountoulakis and Panagiotou show that the push model broadcasts the message to all nodes of a random d regular graph within $(1+o(1))\cdot C_d \cdot \ln n$ rounds, where $C_d = 1/\ln(2(1-1/d)) - 1/(d\cdot\ln(1-1/d))$. In [26], Kempe et al. consider a push type algorithm among uniformly distributed points in R^D and prove that, if each informed node contacts a neighbour in a step according to a certain distribution depending on the distance between the two nodes, then any piece of information is spread to vertices at distance t within $O(\log^{1+\epsilon} t)$ steps, with high probability.
Upper bounds on the running time of the push algorithm in arbitrary networks and especially Cayley graphs are considered in [12]. It is shown that in arbitrary graphs of size n the broadcasting time is bounded, up to a $\log n$ factor, by the mixing time of a corresponding Markov chain. Additionally, a new class of Cayley graphs is introduced on which the push algorithm has optimal performance. Boyd et al. consider the combined push&pull model in arbitrary graphs of size n, and show that the running time is asymptotically bounded by the mixing time of a corresponding Markov chain plus an $O(\log n)$ term [3]. This result is extended to the push model (without allowing pull transmissions) in [35].
In [25], Karp et al. consider the basic random phone call model in complete graphs on n nodes. They present a termination mechanism which, w.h.p., reduces the number of total transmissions to $O(n \log\log n)$, and show that this result is asymptotically optimal. They also consider communication failures and analyse the performance of their method in cases where the connections are established using arbitrary probability distributions. Their results improve a result from [7], which shows a bound of $O(\log n)$ time steps and $O(n\sqrt[3]{\log n})$ message transmissions.
In [11], Elsässer develops an algorithm for the phone call model in random $G_{n,p}$ graphs with $p > \frac{1}{n}\log^{\delta} n$, where $\delta > 2$ is a constant. The algorithm broadcasts a message to all nodes, using time $O(\log n)$ and $O(n(\log\log n + \log n/\log(pn)))$ transmissions, w.h.p. The algorithm has optimal run-time and optimal communication overhead. In [13], the authors consider two simple modifications of the basic random phone call model for $G_{n,p}$ graphs. The first modification allows each node to call *four different* randomly chosen neighbours in every time step, akin to what is phrased *power of multiple choices* elsewhere. The second modification sequentialises this approach and allows the nodes to remember the addresses of the nodes called in the most recent three time steps; these neighbours will not be considered in the current step. Both modifications reduce the number of

transmissions to $O(n \log\log n)$, w.h.p. Furthermore, the idea of avoiding a neighbour contacted in the previous step also leads to sub-logarithmic running time in preferential attachment graphs, see [8]. The proofs use a so called "deconditioning lemma" w.r.t. the distribution of the neighbours of a node after it has randomly chosen $O(\log n)$ neighbours. The proof also integrates some structural properties of $G_{n,p}$ graphs into the dynamical behaviour of randomised broadcasting. However, the deconditioning lemma holds only if the (expected) degree of the graph is large enough (i.e., $\Omega(\log n)$), and the structural integration techniques fail if $d = o(\log n)$. Therefore, it would appear as though the techniques of [13] cannot be generalised to random graphs with sub-logarithmic degrees.
In [9] Doerr et al. propose and analyse a quasi-random analogue to the classical push model where each node has a (cyclic) list of its neighbours, given by an adversary. Once informed, it starts at a random position of the list, but from then on informs its neighbours in the order of the list. For hypercubes or random graphs $G_{n,p}$ they show that $O(\log n)$ rounds suffice to inform every node. These bounds are similar to those in the classical random model. In addition, they prove a $O(\log n)$ bound for sparsely connected random graphs $G_{n,p}$ with $p = (\log n + f(n))/n$, where $f(n) \to \infty$ and $f(n) = O(\log\log n)$. Here, the classical push model needs $\Theta(\log^2(n))$ steps.

1.2 Models and Results

In this paper we consider d-regular random and undirected graphs $G_{n,d}$ on n vertices, with $\delta \leq d \leq \delta \log n$ for some sufficiently large constant δ. We will use (v, w) to name the *undirected* edge between node v and node w. Note that for this choice of d, $G_{n,d}$ is connected w.h.p. [1]. We should also note that the results of [13] can be extended to d-regular random graphs with $d \geq \delta \log n$. Our results can be generalised to a non-regular setting in which the degree of every node is between d and $c \cdot d$ for a constant c. However, for ease of presentation we focus on the truly regular model. We assume that every node knows d, and that it has an estimate of n which is accurate to within a constant factor. We also assume that all nodes have access to a global clock, and that they work synchronously.
In each step every node can create an arbitrary number of messages to be broadcast. Furthermore, in each step every node establishes a channel to four *distinct* neighbours[2]. Once a channel is established between a pair of nodes it may be used bi-directionally. Then, the nodes have to decide which of the established channels to use, and which messages to send over the channel. We assume that they do not know which of their neighbours are aware of a certain message and which are not. We assume that the size of the messages exchanged between a pair of nodes is not limited in any way. The algorithm presented in this paper is *distributed* and *address-oblivious*. An algorithm is called distributed (see [25]) if nodes use only local knowledge to make the decisions as to whether or not to send a message over an open channel. This local information can be e.g. the age and number of broadcast messages they have got, the time the messages arrived, or their own identifier. An algorithm is called address-oblivious if decisions do not depend on the IDs of the nodes which they are connected to via open channels in the current step. Furthermore, in the random phone call model the nodes are not allowed to remember which nodes they communicated with in the previous steps when they choose their next neighbour to exchange messages with (see [25] again). The lower bound we present (see Section 2) unfortunately only applies to a stricter distributed address-oblivious model, where we additionally assume that the decisions only depend on the time

[2]Note that in a sequentialised version of our model, in each step every node v i.u.r. chooses one neighbour from the set of neighbours *not* chosen by v during the last 3 time steps [13]. Clearly, four steps of this sequentialised model can be viewed as one step in the model considered in this paper, and thus, our results can easily be extended to the sequentialised version of our model.

at which the nodes received a message. We refer to this model as *strictly oblivious*. Note that our algorithm also falls into this model. Since in the random phone call model the choices of the communication partners are not allowed to depend on previous steps, we believe that our restricted model is still quite natural (i.e., due to their limited memory, nodes do not remember their previous communication partners, or the history of previous communications).

Our algorithm has time complexity $O(\log n)$ and requires only $O(n \log\log n)$ transmissions per message, a.a.s.[3] In Section 2, we also show a lower bound on the number of transmissions for the standard model of [25]. More precisely, we prove that any distributed strictly oblivious algorithm in the random phone call model needs $\Omega(n \log n/\log d)$ transmissions in order to inform all nodes of a d-regular random graph in an expected time of $O(\log n)$ steps. Our results also imply that the ability to avoid recently chosen neighbours decreases exponentially the expected number of transmissions per message. The authors believe that choosing 3 pairwise distinct neighbours will be sufficient to reduce the number of submissions to $O(n \log\log n)$, but the question is open for two pairwise distinct neighbours.

Several parts of our analysis assume that the input graph is generated by the so-called *configuration model* (also referred to as *pairing model*, see [31] and references therein). In this model, a d-regular random graph is constructed as follows. We start with an empty graph on n nodes, each of which has d *stubs*. In the first step, we i.u.r. choose two stubs and connect the corresponding nodes with an edge. These two stubs are called *matched* thereafter. In each of the next $dn/2 - 1$ steps, we i.u.r. select two unmatched stubs, connect the corresponding nodes with an edge, and these stubs are considered to be matched in any subsequent step.

An alternative description of the process is as follows. We assume the stubs are numbered $1, \ldots, nd$. We pair the first stub with an i.u.r. chosen stub, then we match the next unmatched stub with an i.u.r. chosen unmatched stub, and so on. Note that this process can generate graphs with self-loops and multiple edges with a probability $1 - e^{-O(d^2)}$ [30] (notice that this is the probability of the "bad event"), however, every simple d-regular graph will be generated with equal probability. Also note that it is sufficient to analyse the algorithm for graphs generated with this process (even if the resulting graph is not simple), as long as the failure probability is small enough.

2 Lower Bound

In this section we show the following lower bound for models where each node is allowed to choose one single neighbour in a round. We show that any strictly oblivious, distributed, and time efficient Monte Carlo broadcasting algorithm produces $\Omega(n \log n/\log d)$ message transmissions if, in expectation, less than one node remains uninformed. Recall that a Monte Carlo [28] broadcast algorithm is an algorithm whose running time is upper bounded by some given value, but that might fail (to finish the broadcast) with a certain probability. Note that the algorithms presented in this paper are also Monte-Carlo algorithms with guaranteed running time $O(\log n)$.

Theorem 1. *Let $G_{n,d}$ be a d-regular random graph and assume that a message $\mathcal{M}$ has to be distributed to all nodes of $G_{n,d}$. Let A be a strictly oblivious and distributed broadcast $O(\log n)$-time Monte Carlo algorithm in the random phone call model that finishes the broadcast so that in expectation less than one node remains uninformed. Then A requires at least $\Omega(n \log n/\log d)$ many transmissions of $\mathcal{M}$.*

Proof. Let $G = (V, E)$ be a d-regular random graph. Let $I(t)$ be the set of nodes that are informed by the end of round t, and $H(t)$ the set of uninformed nodes $V \setminus I(t)$ at the end of round t. Assume

[3] A.a.s. means *almost always surely*, i.e., with probability $1 - \log^{-\Omega(1)} n$

the broadcast is finished in time $c \log n$ for some arbitrary positive constant c. We will show that $\epsilon n \log n/(2 \cdot 128^2 c^3 \log d)$ many messages are *not* sufficient, for $\epsilon = 1/(512c^2) + d/n$.
The idea of the proof is as follows. We consider the first time round t in which a constant fraction of the nodes is informed. Note that for any of the informed nodes v and for all rounds $t' \geq t$ we can assume that v's communication pattern in round t' is fixed. Then we show that the informed nodes can not inform all the remaining nodes with the remaining messages.
We assume in the following that at most $\epsilon n \log n/(2 \cdot 128^2 c^3 \log d)$ messages are sent. Let $t = \min\{t' : \ |I(t')| \geq (1-\epsilon)n \text{ and } |H(t')| \leq \epsilon n\}$. Then,

$$(1-\epsilon)n \leq |I(t)| < (1-\epsilon)n + d = (1 - 1/(512c^2))n.$$

For the analysis, we assume that every round is divided into n sub-steps. In sub-step i only the ith node is allowed to send all its messages to its neighbours. Nodes may *not* send messages that they receive in earlier sub-steps of this round. Obviously, this sequentialisation does not change the outcome of the algorithm. We wish to make the presentation of the remainder of the proof as simple as possible, but this requires a minor abuse of notation and model. Specifically, suppose that in this "microscopic" sub-step timing model it is sub-step i of round t that gets the condition on $|I(t)|$ fulfilled for the first time. We now (just as a thought experiment) split round t into two sub-rounds t_1 and t_2, with t_1 comprising sub-steps $1, \ldots, i$, and t_2 comprising sub-steps $i+1, \ldots, n$. For the remainder of the proof we pretend that t_1 and t_2 are complete rounds, so that the order of rounds now is $1 \ldots, t-1, t_1, t_2, t+1, \ldots$ So as not to have to carry these notational anomalies around, we rename the rounds to $0, 1, \ldots, t-1, t, t+1, \ldots$ with (the new) $t-1 = t_1$ and (the new) $t = t_2$.
For our fixed t let $V^{(+)}$ denote the subset of nodes in $I(t)$ that transmit in at least $\log n/(128c \log d)$ many rounds t' with $t' > t$, and let $V^{(-)} = I(t) \setminus V^{(+)}$. By the pigeonhole principle,

$$|V^{(+)}| \leq \frac{\epsilon \cdot n}{2 \cdot 128c^2} = \frac{\epsilon \cdot n}{256c^2}.$$

Let

$$N = \{v \in H(t) : \ v \text{ has fewer than } d/(64c) \text{ neighbours in } V^{(+)}\}$$

Claim 1: $|N| \geq (1 - \frac{1}{4c}) \cdot |H(t)| = (1 - \frac{1}{4c}) \cdot \epsilon n$. Suppose not, that is, $|H(t) \setminus N| \geq \frac{\epsilon n}{4c}$ (the set $H(t) \setminus N$ contains only nodes with fewer than $(1 - \frac{1}{64c})d$ neighbours in $V^{(-)} \cup H(t)$, and thus more than $d/(64c)$ neighbours in $V^{(+)}$). This would imply that these nodes alone were to use more than

$$\frac{\epsilon n}{4c} \cdot \frac{d}{64c} = \frac{\epsilon n d}{256c^2}$$

stubs of nodes in $V^{(+)}$, which contradicts the above upper bound on $|V^{(+)}|$ (there are $d \cdot |V^{(+)}|$ many stubs).

According to the Expander-Mixing Lemma [23], we have

$$\begin{aligned} \left| |E(I(t), H(t))| - \frac{d \cdot |I(t)| \cdot |H(t)|}{n} \right| &= \left| |E(I(t), H(t))| - \frac{d \cdot \epsilon n \cdot (1-\epsilon)n}{n} \right| \\ &\leq 2\sqrt{d-1} \cdot (1 + o(1)) \cdot \sqrt{(1-\epsilon) \cdot n \cdot \epsilon n}, \end{aligned}$$

where we used that the second eigenvalue of a random regular graph is at most $2\sqrt{d-1}(1+o(1))$ with probability at least $1 - 1/n^2$ [18]. Thus, the number of edges between $H(t)$ and $I(t)$ is bounded

from below by

$$\epsilon(1-\epsilon)dn\left(1-\frac{2}{\sqrt{\epsilon(1-\epsilon)d}}\right),$$

with probability $1-n^{-\Omega(1)}$ [18][4]. Then, the number of inner edges in $H(t)$ (meaning edges between nodes of $H(t)$) is upper bounded by $\epsilon^2 nd\cdot(1+\epsilon^2)/2$ if d is large enough. Let

$$N'=\{v\in H(t):\ v \text{ has at most } \tfrac{1}{64c}d \text{ neighbours in } H(t)\ \}.$$

Claim 2: $|N'|\geq(1-\epsilon 128c)\cdot|H(t)|$. Suppose not, that is, $|H(t)\setminus N'|\geq\epsilon^2 128cn$ ($H(t)\setminus N'$ is the set of nodes with more than $\frac{d}{64c}$ neighbours in $H(t)$). This would imply that these nodes alone were to use more than

$$\epsilon^2 128cn\cdot\frac{d}{64c}=2\epsilon^2 nd$$

stubs of nodes in $H(t)$, which contradicts the above upper bound on the number of inner edges in $H(t)$.

Claims 1 and 2 together imply that there are at least

$$\left(1-\frac{1}{4c}-\epsilon 128c\right)\cdot|H(t)| \tag{1}$$

nodes $S\subset H(t)$ having at most $d/(64c)$ neighbours in $V^{(+)}$ or in $H(t)$, w.h.p. Applying Equation (1) yields

$$|S|\geq\left(1-\frac{1}{4c}-\epsilon 128c\right)\cdot|H(t)|>|H(t)|/2=\epsilon n/2.$$

We show that S contains a matching of size $\epsilon^2 n/9$. We apply the Expander-Mixing Lemma to the set of edges $E(S,\overline{S})$ between S and $V\setminus S$ and obtain that the number of inner edges in S is lower bounded by $\epsilon^2 nd/4\cdot(1-\epsilon^2)$, whenever d is large enough. We compute a matching by repeatedly removing arbitrary edges (and adding them to our matching) as well as all at most $2d-2$ edges incident to either endpoint.

Of the $\epsilon^2 n/9$ many pairs, each of the two nodes has at least $(1-\frac{1}{32c})d$ many neighbours in $V^{(-)}$. Let $\mathcal{P}'$ denote those nodes. Fix $v\in\mathcal{P}'$, and consider table $(a^v)_{t,i}$ with $1\leq t\leq c\log n$ and $1\leq i\leq(1-\frac{1}{32c})d$, that is, we have one row for each time step, and one column for each of v's first $(1-\frac{1}{32c})d$ many neighbours in $V^{(-)}$. Let $a^v_{t,i}=1$ if v's i-th neighbour in $V^{(-)}$ is quiet in step t, and $a^v_{t,i}=0$ otherwise.

Since those neighbours are in $V^{(-)}$, they transmit in fewer than $\log n/(128c\log d)$ many steps, that is, are quiet in at least

$$c\log n-\frac{\log n}{128c\log d}=\left(c-\frac{1}{128c\log d}\right)\cdot\log n$$

many steps. Therefore, each column in our table has at least $(c-1/(128c\log d))\cdot\log n$ many 1 entries, and the total number of 1 entries is at least

$$T\stackrel{\text{def}}{=}\left(1-\frac{1}{32c}\right)\cdot d\cdot\left(c-\frac{1}{128c\log d}\right)\cdot\log n. \tag{2}$$

[4]To apply this result for simple graphs, the degree should be independent of n. For higher degrees, some more probabilistic analysis is needed which is omitted in the lower bound case.

Now we use yet another pigeonhole argument to show that most of the rows are mostly filled with 1 entries. These steps are called *quiet* steps in the following.

Claim 3: At least $(c - \frac{1}{4\log d})\log n$ of the rows are quiet, i.e. have at least a $(1 - \frac{1}{32c})$-fraction of 1 entries. Suppose this was not true, i.e., fewer than $(c - \frac{1}{4\log d})\log n$ rows had at least $(1 - \frac{1}{32c})^2 d$ many 1 entries, i.e., a $(1 - \frac{1}{32c})$-fraction of the width of the table, which itself is $(1 - \frac{1}{32c})d$. In this case, there would be strictly fewer than

$$\begin{aligned}
&\overbrace{\left(c - \frac{1}{4\log d}\right)\log n \cdot \left(1 - \frac{1}{32c}\right) d}^{\text{rows assumed full}} + \overbrace{\left(\frac{\log n}{4\log d}\right) \cdot \left(1 - \frac{1}{32c}\right)^2 d}^{\text{remaining rows}} \\
&= \left(1 - \frac{1}{32c}\right) d \log n \cdot \left[\left(c - \frac{1}{4\log d}\right) + \left(\frac{1 - \frac{1}{32c}}{4\log d}\right)\right] \\
&= \left(1 - \frac{1}{32c}\right) d \log n \cdot \left(c - \frac{1}{128c\log d}\right) = T
\end{aligned}$$

in total, contradicting bound (2).

Our goal now is to show that the expected number of $\mathcal{P}$-pairs that do not receive the message is larger than one. Let (u, v) be an arbitrary $\mathcal{P}$-pair that has both nodes in S. In the following we consider first pull transmissions and then push transmissions.

Pull: In the following we calculate a lower bound p_u on the probability that u opens a channel to v in the at most $\log n/(4\log d)$ non-quiet steps and that in the other steps u opens channels only to neighbours in $V^{(-)}$, which do not transmit in these steps. The bound p_v is defined accordingly. The probability that u opens a channel to v in the non-quiet steps can be lower bounded by $d^{-\log n/(4\log d)}$. In the quiet steps u is only to open channels to neighbours in $V^{(-)}$ that do not transmit. This probability can be lower bounded by $(1 - 1/(32c))^{2(c-1/(4\log d))\log n}$. Observe that $(1 - \frac{1}{z})^z \geq \frac{1}{4}$ for $z \geq 2$. Then

$$\begin{aligned}
\left(\frac{1}{d}\right)^{\frac{\log n}{4\log d}} \cdot \left[\left(1 - \frac{1}{32c}\right)^2\right]^{(c - \frac{1}{4\log d})\log n} &= \left(\frac{1}{d}\right)^{\frac{\log n}{4\log d}} \cdot \left(1 - \frac{1}{32c}\right)^{2(c - \frac{1}{4\log d})\log n} \\
&= \left(\frac{1}{n}\right)^{\frac{1}{4}} \cdot \left(1 - \frac{1}{32c}\right)^{32c\left(\frac{(c - \frac{1}{4\log d})}{16c}\log n\right)} \\
&> \left(\frac{1}{n}\right)^{\frac{1}{4}} \cdot \left(\frac{1}{4}\right)^{\frac{c - \frac{1}{4\log d}}{16c}\log n} \\
&= \left(\frac{1}{n}\right)^{\frac{1}{4}} \cdot \left(\frac{1}{n}\right)^{2\frac{c - \frac{1}{4\log d}}{16c}} \\
&= \left(\frac{1}{n}\right)^{\frac{1}{4}} \cdot \left(\frac{1}{n}\right)^{\frac{1}{8} - \frac{1}{32c\log d}} \\
&= \left(\frac{1}{n}\right)^{\frac{3}{8} - \frac{1}{32c\log d}} =: p_u.
\end{aligned}$$

Of course, the same probability bounds hold for node v. Thus, the probability that both u and v do only communicate with each other in non-quiet steps and with nodes that do not transmit in quiet steps can be lower bounded by $p_u \cdot p_v$.

Push: Now we calculate q_v as a bound on the probability that v does not receive the message due to a push transmission. The value q_v covers the cases that v neither receives a `push(`$\mathcal{M}$`)` transmission from a neighbour in $V^{(+)}$ (or from an informed node of $H(t)$), nor a node $w \in V^{(-)}$ directs a channel to v in a step in which w transmits a message. The probability of the first event can be lower bounded by $(1-1/d)^{c\log n \cdot d/32c} = (1-1/d)^{d\log n/32}$. The probability of the second event can be lower bounded by $(1-1/d)^{d\log n/(128c\log d)}$. Hence

$$\begin{aligned}
\left(1-\frac{1}{d}\right)^{\frac{d\log n}{32}+\frac{d\log n}{128c\log d}} &= \left(1-\frac{1}{d}\right)^{d\log(n)\cdot\left(\frac{1}{32}+\frac{1}{128c\log d}\right)} \\
&> \left(\frac{1}{4}\right)^{\log(n)\cdot\left(\frac{1}{32}+\frac{1}{128cd\log d}\right)} \\
&= \left(\frac{1}{n}\right)^{2\cdot(\frac{1}{32}+\frac{1}{128cd\log d})} \\
&> \left(\frac{1}{n}\right)^{\frac{1}{16}+\frac{1}{64cd\log d}} =: q_v.
\end{aligned}$$

Now we calculate a lower bound q'_u on the probability that u does not receive the message due to a push transmission, under the condition that v neither receives a `push(`$\mathcal{M}$`)` transmission from a neighbour in $V^{(+)}$ (or in $H(t)$), nor that a node $w \in V^{(-)}$ directs a channel to v in a step in which w transmits a message. Similar to the calculation above we get

$$\left(1-\frac{1}{d-1}\right)^{\frac{d\log n}{32}+\frac{d\log n}{128c\log d}} > \left(\frac{1}{n}\right)^{\frac{1}{16}+\frac{1}{64cd\log d}} =: q'_u.$$

The probability that u and v do not receive the message at all is at least

$$\begin{aligned}
p_u \cdot p_v \cdot q'_u \cdot q_v &= \left[\left(\frac{1}{n}\right)^{\frac{3}{8}-\frac{1}{32c\log d}}\right]^2 \cdot \left[\left(\frac{1}{n}\right)^{\frac{1}{16}+\frac{1}{64cd\log d}}\right]^2 \\
&= \left[\left(\frac{1}{n}\right)^{\frac{3}{4}-\frac{1}{16c\log d}}\right] \cdot \left[\left(\frac{1}{n}\right)^{\frac{1}{8}+\frac{1}{32cd\log d}}\right] \\
&= \left(\frac{1}{n}\right)^{\frac{7}{8}-\frac{1}{16c\log d}+\frac{1}{32cd\log d}} > \left(\frac{1}{n}\right)^{8/9}
\end{aligned}$$

for d large enough. Since $\mathcal{P}' \geq \epsilon^2 n/9$, the expected number of uninformed pairs is at least

$$\left(\frac{\epsilon^2 n}{9}\right) \cdot n^{-8/9} = \left(\frac{\epsilon^2}{9}\right) \cdot n^{1/9} > 1$$

for n large enough. Hence, we can assume that the expected number of uninformed nodes is at least two. □

3 Broadcasting Algorithm

In the modified model described in the introduction, in each step every node chooses four distinct neighbours instead of one. In each time step, whenever a communication channel is established between two nodes, each of them has to decide which messages to transmit, or whether to transmit a message at all, without knowing if the node at the other end of the edge has already received a given message or not. In other words, *opening* a channel does not, in general, imply *transmission* of a message. In the following we define some procedures which are frequently used by each node of the graph.

`open:` Choose four *distinct* neighbours uniformly at random and establish communication channels to them. These channels are called *outgoing* in the following. The procedure also establishes communication channels with all nodes which call the corresponding node. Those channels are called *incoming*.

`push(`$\mathcal{M}$`):` Send message $\mathcal{M}$ over all outgoing channels.

`pull(`$\mathcal{M}$`):` Send message $\mathcal{M}$ over all incoming channels.

`receive:` Receive and store all messages coming over open channels in $\mathcal{M}$ (if any).

`close:` Close all channels opened in the current round.

In each step t, every node $u \in V$ executes the procedure given in Algorithm 1 or Algorithm 2, depending on the degree of the graph (which we assumed the nodes to be aware of). *The algorithm will be run for every message.* The nodes decide if a message has to be transmitted via push or pull, depending on the time at which the message has been generated. When more than one message is considered, the node combines to a single message all messages which should be transmitted via push (pull), and forwards this combined message over all open outgoing (incoming) channels. In the following we state the algorithm (w.l.o.g.) for one fixed message $\mathcal{M}$ and we assume that the message is created in time step 0. Hence, the age of the message is nothing else than the current time step (both denoted by t).
Algorithm 1 depicts the algorithm for $\delta \leq d \leq \delta \log\log n$.
In Algorithm 1 the parameter α is a sufficiently large constant, and the nodes are initially *not* in state `active`. The algorithm has four distinct phases, and each of these phases comprises several steps. In the first phase every informed node transmits the message exactly once. At the end of the phase a constant fraction of the nodes are informed. In the second phase the informed nodes perform push transmissions during all steps of this phase. At the end of the phase we have at most $O(n/\log^5 n)$ uninformed nodes. The third phase consists only of one step in which every informed node sends out the message over all incoming channels (`pull(`$\mathcal{M}$`)`). All nodes which received the message for the first time in this step will be in state `active`. In the last phase, in each step all newly informed nodes become `active` as well, and all active nodes transmit the message via `push(`$\mathcal{M}$`)` during all subsequent steps of this phase.

For $\delta \log\log n \leq d \leq \delta \log n$ we have to modify the algorithm slightly. We refer to the modified version as Algorithm 2. In Algorithm 2, Step 6 (Phase 3) is replaced by "**if** $\lceil \alpha \log n + \log\log n \rceil + 1 \leq t \leq \lceil \alpha \log n + 2\alpha \log\log n \rceil$ **then**", and Steps 8, 9 and 10 are removed.

Algorithm 1

1: `open`

2: **if** $t \leq \lceil \alpha \log n \rceil$ **then** {Phase 1}

3: **if** the message is created or received for the first time in the previous step **then** `push(`$\mathcal{M}$`)`.

4: **if** $\lceil \alpha \log n \rceil + 1 \leq t \leq \lceil \alpha(\log n + \log\log n) \rceil$ **then** {Phase 2}

5: **if** the node is informed **then** `push(`$\mathcal{M}$`)`.

6: **if** $t = \lceil \alpha(\log n + \log\log n) \rceil + 1$ **then** {Phase 3}

7: **if** the node is informed **then** `pull(`$\mathcal{M}$`)`.

8: **if** $\lceil \alpha(\log n + \log\log n) \rceil + 2 \leq t \leq 2 \cdot \lceil \alpha \log n \rceil + \lceil \alpha \log\log n \rceil$ **then** {Phase 4}

9: **if** the message is received for the first time in the previous step (Phase 3 or 4) **then** go to state `active`.

10: **if** `active` **then** `push(`$\mathcal{M}$`)`.

11: `receive`

12: `close`

Algorithm 2

1: `open`

2: **if** $t \leq \lceil \alpha \log n \rceil$ **then** {Phase 1}

3: **if** the message is created or received for the first time in the previous step **then** `push(`$\mathcal{M}$`)`.

4: **if** $\lceil \alpha \log n \rceil + 1 \leq t \leq \lceil \alpha(\log n + \log\log n) \rceil$ **then** {Phase 2}

5: **if** the node is informed **then** `push(`$\mathcal{M}$`)`.

6: **if** $\lceil \alpha(\log n + \log\log n) \rceil + 1 \leq t \leq \lceil \alpha \log n + 2\alpha \log\log n \rceil$ **then** {Phase 3}

7: **if** the node is informed **then** `pull(`$\mathcal{M}$`)`.

8: `receive`

9: `close`

4 Analysis of the Algorithm

In this section we analyse the behaviour of Algorithm 1 on G. We assume that $\alpha \log n$ and $\alpha \log\log n$ are integers, and hence $\lceil \alpha \log n \rceil = \alpha \log n$ and $\lceil \alpha \log\log n \rceil = \alpha \log\log n$.
Note that Phase 1 and Phase 2 are the same for both Algorithm 1 and Algorithm 2. In Section 4.1 we analyse Phase 1 and show that by the end of Phase 1 a constant fraction of the nodes is informed. In Section 4.2 (Lemma 3) we analyse Phase 2 and show that by the end of the phase there are at most $O(n/\log^5 n)$ uninformed nodes. In Section 4.3 we analyse the remaining phases and show our main results for large and small degrees.
We first need a few more definitions.

- $I(t)$ is the set of informed nodes after step t but before step $t+1$.
- Let $I^+(t) = I(t) \setminus I(t-1)$, that is, the nodes that become informed in step t.
- $H(t)$ is the set of uninformed nodes $V \setminus I(t)$ at the end of step t, with $h(t) = |H(t)|$.
- We call an edge *used before step t* if one of the nodes incident to this edge transmitted the information along this edge before step t. Let $U(t) \subseteq V$ be the set of nodes incident to at least one edge which is not used before step $t+1$. Notice that $H(t) \subseteq U(t)$.
- $I_C(t) = \{v \in I(t) \mid \exists u \in H(t) : \ (u,v) \in E\}$ is the set of informed nodes connected to an uninformed node.
- $H_C(t) = \{u \in H(t) \mid \exists v \in I(t) : \ (u,v) \in E\}$ is the set of uninformed nodes connected to an informed one.
- $E(S, \overline{S})$ is the set of edges between S and $\overline{S}$, where $S, \overline{S} \subset V$.
- For $1 \le i \le 5$, let $H_i(t)$ denote the set of nodes in $H(t)$ with at least i neighbours in $H(t)$ at the beginning of round t, with $h_i(t) = |H_i(t)|$.

4.1 Phase 1

In the next two lemmas we prove that within the first $\alpha \log n$ steps at least $n/8$ nodes of $G_{n,d}$ become informed, w.h.p., whenever α is large enough.
Recall that Phase 1 consists of $\alpha \log n$ many steps. Every node that receives the message for the first time in step t performs a push transmission in the next step. In Lemma 1 we show that w.h.p. the set of informed nodes grows by a constant factor in each of the first $\alpha \log\log n$ steps. Lemma 2 shows the same for the remaining steps of the phase.

Lemma 1 (Phase 1)**.** *Assume $t < \alpha \log\log n$.*

1. *If $d < \sqrt[3]{\log n}$ then we have w.h.p. that $|I^+(t+1)| \ > \ 2 \cdot |I^+(t)|$.*
2. *If $d \ge \sqrt[3]{\log n}$ and $|I^+(t)| \le d/\sqrt[10]{\log n}$, then a.a.s. we have $|I^+(t+1)| \ > \ 2 \cdot |I^+(t)|$.*

Proof. First of all, note that for $t < \alpha \log n$ all newly informed nodes perform exactly one push transmission in the first phase of the algorithm.

Case 1: $d < \sqrt[3]{\log n}$**.** We shall apply the principle of deferred decisions: we pretend that the random choices of our algorithm actually *construct* (part of) the graph $G_{n,d}$ under consideration. At time 0 we are given n nodes with d unmatched stubs each. In round $t+1$, each *newly informed node*, that is, each $v \in I^+(t)$, selects four of its (matched or unmatched) stubs i.u.r. without

replacement. For each vertex, every unmatched among the selected four stubs will now be matched (connected) with one i.u.r. chosen unmatched stub. Notice that there is always a sufficient number of stubs available as we allow for two nodes in $I^+(t)$ to match their stubs selected in this step. In the following we call the set of edges created in the first t steps E^t.

To prove the result we show by induction that the following holds with a probability of $1 - \log^{O(1)}(n)/n$:

1. None of the nodes of $I^+(t+1)$ are incident to more than one edge in E^{t+1}.
2. $|I^+(t+1)| > 2 \cdot |I^+(t)|$.

For $t = 0$, the probability that two of the four edges chosen by the only node in $I(0)$ are directed to the same node $u \in V$ is at most

$$\binom{4}{2} \cdot \frac{4d}{dn} = \frac{6}{n}.$$

Hence, with a probability of at least $1 - \log^{O(1)}(n)/n$ there will be at least two nodes in $I(1)$.

For $t \geq 1$ we can assume (due to the induction hypothesis) that with a probability of $1-\log^{O(1)}(n)/n$ we have $|I^+(t)| > |I^+(t-1)|$, and none of the nodes of $I^+(t)$ has more than one neighbour in $I^+(t-1)$. Now we show that, under this condition, the invariant is still fulfilled for step t. The lemma then follows using the union bound over all steps of Phase 1.

Recall that $I^+(t)$ is the set of nodes that became informed during step t by some node(s) of $I^+(t-1)$, i.e., they were not informed before but are informed after step t – the construction is iterative, and only nodes newly informed in some step pick other nodes in the next step. (It may be helpful to visualise the process in terms of layers of an onion.) Recall that in every step of the algorithm every node opens a channel to four distinct (randomly chosen) of its d many neighbours. Then, due to the induction hypothesis, each node chooses at least three unmatched stubs. Now, divide step $t+1$ into $4|I^+(t)|$ sub-steps. Let $u_1, \ldots, u_{|I^+(t)|}$ be the elements of $I^+(t)$ ordered in this way. In sub-step $4(t'-1)+i+1$, $0 \leq i \leq 3$, node $u_{t'}$ chooses one of its stubs, and pairs this stub (if not matched yet) with some unmatched stub in the graph. Since $t \leq \alpha \log\log n$ we have that $|I(t)| = \log^{O(1)} n$. Thus, if the stub chosen by node $u_{t'}$ in sub-step $4(t'-1)+i+1$ is itself unmatched, then this stub is paired with the stub of a node which has only unmatched stubs at this time with probability larger than

$$\frac{dn - (d \cdot |I(t)| + d \cdot (4(t'-1)+i+1))}{dn} \geq 1 - \log^{O(1)}(n)/n,$$

regardless of the matchings established so far. dn is the total number of stubs of all the nodes in G. From that number we subtract all stubs to nodes in $I(t)$ as well as the stubs to the nodes $u_1, \ldots, u_{t'}$. Thus, a node of $I^+(t+1)$ has more than one neighbour in $I^+(t)$ with probability smaller than $\log^{O(1)} n/n$. Since $|I^+(t+1)| = \log^{O(1)} n$, the union bound implies that, given the induction hypothesis holds for $t-1$, there is no node in $|I^+(t+1)|$ with more than one neighbour in $I^+(t)$ with conditional probability at least $1 - \log^{O(1)}(n)/n$. This shows the first invariant. Moreover, each node of $I^+(t)$ has at least 3 neighbours in $I^+(t+1)$ with the same conditional probability. Hence, the second invariant is maintained, too.

Recall that this pairing process generates graphs with self-loops and multiple edges with a probability $1 - e^{-O(d^2)}$ (see [30]). (This bound on the simplicity of the generated graph is precisely what requires us to follow a different approach for larger degrees – recall that the $1 - e^{-O(d^2)}$ is the probability for the bad event happening.) Since $t = O(\log\log n)$ we can assume that the probabilities calculated above also hold for standard random d-regular graphs without self loops.

Case 2: $d \geq \sqrt[3]{\log n}$**.** Again, we show the claim by induction. However, this time we assume that a simple graph is chosen uniformly at random from the space of all d-regular graphs. Since $|I^+(t)| \leq d/\sqrt[10]{\log n}$, and $2 \cdot |I^+(t'-1)| \leq |I^+(t')|$ for any $t' \leq t$, each node of $I^+(t)$ can have at most $O(d/\sqrt[10]{\log n})$ neighbours in $I(t+1)$.
Let step t be divided into $|I^+(t)|$ many sub-steps. Let $u_1, \ldots, u_{|I^+(t)|}$ be the nodes of $I^+(t)$. We perform our induction over the sub-steps. In sub-step i, node u_i chooses four distinct neighbours and sends a copy of the information to these neighbours as described in Algorithm 1. Let $H_i(t)$ be the set of uninformed nodes in sub-step i of step t. Then, in sub-step i a node u_i chooses neighbours exclusively from $H_{i-1}(t)$ with probability at least $(1-O(1/\sqrt[10]{\log n}))^4 = 1-O(1/\sqrt[10]{\log n})$, regardless of the push transmissions performed by the nodes of $I(t-1) \cup \{u_1, \ldots, u_{i-1}\}$. Define p as the probability that in sub-step i a node chooses neighbours exclusively from $H_{i-1}(t)$. Then $p = O(1/\sqrt[10]{\log n})$ and

$$\Pr[|I^+(t+1)| < 2|I^+(t)|] \leq \sum_{i=\frac{|I^+(t)|}{2}}^{|I^+(t)|} \binom{|I^+(t)|}{i} \cdot p^i \cdot (1-p)^{|I^+(t)|-i}.$$

To bound this binomial distribution we apply Lemma 9 (in the appendix) and obtain that

$$\Pr[|I^+(t+1)| < 2|I^+(t)|] \leq (2p)^{\frac{|I^+(t)|}{2}} \cdot (2 \cdot (1-p))^{\frac{|I^+(t)|}{2}} = \frac{1}{\log^{\Omega(1)} n}.$$

Thus, $|I^+(t+1)| \geq 2 \cdot |I^+(t)|$ a.a.s. □

According to Lemma 1 the number of newly informed nodes in step $t = \alpha \log\log n$ becomes larger than $\log^q n$ for $d < \sqrt[3]{\log n}$ and an appropriate constant q. On the other side, the amount of newly informed nodes becomes at least $d/\sqrt[10]{\log n}$ for $d \geq \sqrt[3]{\log n}$. In the following lemma we consider two cases separately:

1. $d \geq \sqrt[3]{\log n}$ and $|I^+(t)| \in \{d/\sqrt[10]{\log n}, \ldots, \log^q n\}$
2. $|I^+(t)| \geq \log^q n$

We show by induction that, for $d \geq \sqrt[3]{\log n}$, within the next $2\alpha \log\log n$ steps we have a.a.s. that $|I^+(t)| \geq \log^q n$, whenever α is large enough. Then we consider the case $|I^+(t)| \geq \log^q n$ for arbitrary values of d.

Lemma 2 (Phase 1)**.** *Assume $\alpha \log\log n \leq t < \alpha \log n$. There exists a constant $c > 1$ such that the following hold.*

1. *Assume $d \geq \sqrt[3]{\log n}$ and $\log^q n \geq |I^+(t)| > d/\sqrt[10]{\log n}$. Then $|I^+(t+1)| > c \cdot |I^+(t)|$ with probability $1 - e^{-\log^{\Omega(1)} n}$.*
2. *Assume $\log^q n \leq |I^+(t)| \leq n/8$, where q is a (large) constant. Then w.h.p. $|I^+(t+1)| > c \cdot |I^+(t)|$.*

Proof. We first consider small values of $I^+(t)$.

1. In step $t+1$ each vertex of $I^+(t)$ chooses four of its stubs. The unmatched but chosen stubs are now matched with some other i.u.r. chosen unmatched stubs. Let $u_1, \ldots, u_{|I^+(t)|}$ be the nodes in $I^+(t)$. We divide step $t+1$ into $4|I^+(t)|$ sub-steps. In sub-step $4(i-1)+j+1$ we connect the $(j+1)$-st stub of node u_i, where $i \leq |I^+(t)|$ and $0 \leq j \leq 3$.

   We denote by $I^+_{>20}(t)$ and $I^+_{\leq 20}(t)$ the subsets of nodes of $I^+(t)$ with more than 20 and at most 20 matched stubs, respectively. We first show that there are at most $|I^+(t)|/5$ nodes with more than 20 matched stubs by the end of step t.

   Using the pigeonhole principle it is easy to argue there are at most $|I^+(t)|/5$ nodes in $I^+(t)$ with at least 20 matched stubs by the end of step t. If there were more, this would imply

$$4|I^+(t-1)| \geq \frac{|I^+(t)|}{5} \cdot 20,$$

   equivalent to $|I^+(t-1)| \geq |I^+(t)|$. Then we have $I^+_{>20}(t) \leq |I^+(t)|/5$ and

$$I^+_{\leq 20}(t) \geq \frac{4 \cdot |I^+(t)|}{5} \geq \frac{4d}{5\sqrt[10]{\log n}}.$$

   W.l.o.g. let $u_1, \ldots, u_{|I^+_{\leq 20}(t)|}$ be the elements of $I^+_{\leq 20}(t)$. Let $H_i(t)$ be the set of uninformed nodes in step $t+1$ after sub-step i. According to the pairing model, u_{i+1} has at most $4/5$ of its neighbours in $H_i(t)$ (or at least $1/5$th of the neighbours in $V \setminus H_i(t)$) with probability upper bounded by

$$\sum_{i=\frac{4d}{5}}^{d-20} \binom{d-20}{i} p^i (1-p)^{d-20-i} \leq \left(\frac{5p}{4}\right)^{\frac{4d}{5}} (5(1-p))^{\frac{d}{5}} \leq \left(\frac{10p}{4}\right)^{\frac{4d}{5}} = \left(\frac{\log^{O(1)} n}{n}\right)^{\frac{4d}{5}},$$

   where $p = \log^{O(1)}(n)/n$. This holds independently of the push transmissions performed in steps $1, \ldots, t$ and sub-steps $1, \ldots, 4i$ of step $t+1$.

   We define random variables X_i, $1 \leq i \leq I^+_{\leq 20}(t)$, with $X_i = 1$ if u_i has at most $4/5$th of its neighbours in $H_i(t)$ and $X_i = 0$ otherwise. For

$$X = \sum_{i=1}^{|I^+_{\leq 20}(t)|} X_i$$

   we have

$$\mathbb{E}[X] = |I^+_{\leq 20}(t)| \cdot \left(\frac{\log^{O(1)} n}{n}\right)^{4d/5} \geq \frac{4d}{5\sqrt[10]{\log n}} \cdot \left(\frac{\log^{O(1)} n}{n}\right)^{4d/5}.$$

   To bound the tails of X we apply the Chernoff bound from Lemma 10 in the appendix and obtain

$$\Pr\left[X \geq \frac{1}{5} \cdot |I^+_{\leq 20}(t)|\right] \leq \left(5e\left(\frac{\log^{O(1)} n}{n}\right)^d\right)^{|I^+_{\leq 20}(t)|/5} = \left(\frac{\log^{O(1)} n}{n}\right)^{\Theta(d|I^+_{\leq 20}(t)|)} = 1/e^{d^{2+\Omega(1)}}.$$

Hence, with a probability of $1/e^{d^{2+\Omega(1)}}$ we have that 4/5 of the nodes of $I^+_{\leq 20}(t)$ will have more than 4/5 of their neighbours in $H_i(t)$. Let us refer to these nodes as $J^+_{\leq 20}(t)$ in the following, and assume $|J^+_{\leq 20}(t)| \geq 4/5 \cdot |I^+_{\leq 20}(t)|$. The probability that in step $t+1$ one of the nodes in $J^+_{\leq 20}(t)$ chooses four distinct neighbours in $H_i(t)$ is $(4/5)^4$. This probability is independent for different nodes in $J^+_{\leq 20}(t)$. Thus, the Chernoff bound from Lemma 10 implies that the probability that a fraction of more than 2/3 of the nodes of $J^+_{\leq 20}(t)$ chooses fewer than 4 distinct neighbours from $H_i(t)$ is

$$\exp(-\Theta(|J^+_{\leq 20}(t)|)) \leq \exp(-\Theta(|I^+_{\leq 20}(t)|)).$$

Hence,

$$\Pr[|I^+(t+1)| \geq 4 \cdot \frac{4}{5} \cdot \frac{1}{3} \cdot |I^+_{\leq 20}(t)|] \geq 1 - e^{-\Theta(I^+_{\leq 20}(t))} - 1/e^{d^{2+\Omega(1)}} = 1 - 1/e^{\Theta(|I^+(t)|)},$$

and the claim follows.

2. The claim is shown using an induction. Recall that we inherited the case distinction from Lemma 1. As in the proof of Lemma 1, we base our analysis on the so-called *configuration model* for generating random regular graphs and apply the principle of deferred decisions. Again, we start with an empty graph on n nodes and d *stubs* per node. In round $t+1$, each newly informed node selects four of its stubs i.u.r. without replacement and these stubs will now be matched (connected) with i.u.r. chosen unmatched stubs.

   First we state the following technical statement.

   Claim 1. *Assume $t < \alpha \log n$ and $|I(t)| \leq n/8$. With probability $1 - e^{-\omega(\log^3 n)}$, the number of edges generated so far between $I^+(t)$ and $H(t)$ is at least $(83/40) \cdot |I^+(t)|$.*

   Proof. Recall that a stub s of a node $v \in I^+(t)$ is called *matched* if s was paired with a stub chosen by a node in $I^+(t-1)$ in step $t-1$. Notice that step $t-1$ is the only possibility as otherwise $v \notin I^+(t)$. Due to the induction hypothesis we may assume $|I^+(t)| > |I^+(t-1)|$ and that $I^+(t)$ is the set of nodes that become informed during step t by some node(s) of $I^+(t-1)$.

   Recall that there are at most $|I^+(t)|/5$ nodes in $I^+(t)$ with at least 20 matched stubs by the end of step t.

   Now back to the algorithm (as a side note, it may be helpful to consider the following as some sort of "fusion" of graph-generation process and broadcasting algorithm, with the focus being on making both "match"). Recall that a node chooses four out of its d many edges. For nodes $v \in I^+_{\leq 20}(t)$, with probability at least $(1 - 20/(d-2))^3$ at least three out of these four edges correspond to unmatched stubs. Now we define random variables X_i, $1 \leq i \leq I^+_{\leq 20}(t)$ with $X_i = 1$ if the ith node of $I^+_{\leq 20}(t)$ chooses at least 3 of its unmatched stubs and $X_i = 0$ otherwise. For

$$X = \sum_{i=1}^{|I^+_{\leq 20}(t)|} X_i$$

we have

$$\mathbb{E}[X] = |I^+_{\leq 20}(t)| \cdot (1 - 20/(d-2))^3.$$

Since the choices of different nodes are independent we may apply the Chernoff bound from Lemma 10 for bounding the tails of X.

$$\Pr[X \leq (1 - o(1)) \cdot \mathbb{E}[X]] \leq e^{-(1-o(1))^2 \cdot |I^+_{\leq 20}(t)| \cdot (1-20/(d-2)^3)/2}.$$

Hence, at least $|I^+_{\leq 20}(t)| \cdot (1-20/(d-2))^3 \cdot (1-o(1))$ nodes will choose at least three unmatched stubs with probability

$$1 - e^{-o(1) \cdot |I^+_{\leq 20}(t)| \cdot (1-20/(d-2)^3)} = 1 - o(e^{-\log^3 n}).$$

The last equality holds since $|I^+_{\leq 20}(t)| \geq 4 \log^q n/5$. Thus, if Y denotes the number of edges between $I^+(t)$ and $H(t)$, then

$$\begin{aligned}
\mathbb{E}[Y] &\geq 3 \cdot \frac{4}{5} \cdot |I^+(t)| \cdot \left(1 - \frac{20}{d-2}\right)^3 \cdot (1 - o(1)) \cdot \frac{|H(t)| \cdot d - 2}{nd - 2} \\
&\geq 3 \cdot \frac{4}{5} \cdot |I^+(t)| \cdot \left(1 - \frac{20}{d-2}\right)^3 \cdot (1 - o(1)) \cdot \frac{7n/8 \cdot d - 2}{nd - 2} \\
&\geq |I^+(t)| \cdot \left(1 - \frac{20}{d-2}\right)^3 \cdot \left(\frac{84}{40} \cdot (1 - o(1))\right) \\
&= |I^+(t)| \cdot \left(\frac{84}{40} \cdot (1 - o(1))\right),
\end{aligned}$$

which is larger than $83.5 \cdot |I^+(t)|/40$ whenever d is large enough.

To conclude the proof, let $I^+_{\leq 20}(t) = \{v_1, \ldots, v_k\}$. Let G'_0 be the graph with nodes $\{v_1, \ldots, v_k\} \cup H(t)$ and no edges. Let G'_i be the graph with node set $\{v_1, \ldots, v_k\} \cup H(t)$ that contains all edges from G_i and an edge from v_i to $u \in H(t)$ if v_i has an edge to node u. For every $i \in [k]$ let $Y_i = \mathbb{E}[Y \mid G'_i]$. Then $\mathbb{E}[Y_i \mid Y_0, \ldots, Y_{i-1}] = Y_{i-1}$ for any $i \geq 0$, and $(Y_i)_{i \in [k]}$ is a Martingale sequence. Since every node v_i has at most 4 edges, the Martingale sequence satisfies the bounded difference condition with value 4. We can apply the Azuma-Hoeffding inequality (see Lemma 11 in the appendix). Hence,

$$\Pr[|Y - \mathbb{E}[Y]| \leq (0.5/80) \cdot |I^+(t)|] \leq e^{-\omega(\log^3 n)}.$$

This finishes the proof of the claim. □

To continue the proof of the lemma we show that at least half of these edges between $I^+(t)$ and $H(t)$ are connected to distinct vertices of $H(t)$. Let $\ell = 83|I^+(t)|/40$. Let $\{s_i \mid i \in [\ell]\}$ represent the first ℓ stubs in $I^+(t)$ which are connected to some vertices of $H(t)$. Again, we define a graph by adding the edges (defined by these stubs) one by one. Let G''_0 be the graph with the nodes $\{s_1, \ldots, s_\ell\} \cup H(t)$ and no edges. Let G''_i be the graph that contains all edges from G''_{i-1} and the edge defined by the ith stub. Let Y be the number of vertices of $H(t)$ that

are connected to at least one of the stubs $\{s_1, \ldots, s_\ell\}$. Clearly, the ℓth stub is connected to a node in $H(t)$, which is not connected to any stub s_j with $j \in \{1, \ldots, \ell-1\}$, with probability at least $(1 - d/(|H(t)| \cdot d - (\ell-1))^{\ell-1}$. Here, $d/(|H(t)|d - (\ell-1))$ represents an upper bound on the probability that the ℓ-th stub connects to a node, which has already been matched to some s_j. The $\ell - 1$ in the exponent is an upper bound on the number of nodes already matched with some previous stub s_j. Then

$$\mathbb{E}[Y] \geq \ell \cdot \left(1 - \frac{d}{|H(t)| \cdot d - (\ell-1)}\right)^{\ell-1} > \frac{\ell}{2} = \frac{83}{80} \cdot |I^+(t)|.$$

We again define a sequence $Y_i = \mathbb{E}[Y \mid G_i'']$. It is easy to verify that Y_i is a Martingale sequence. Since every node v_i has at most 4 edges, the Martingale sequence satisfies the bounded difference condition with value 4. We can apply the Azuma-Hoeffding inequality (Lemma 11) to obtain

$$\Pr[|Y - \mathbb{E}[Y]| \geq (1/80) \cdot |I^+(t)|] \leq e^{-\log^{\Omega(1)} n}.$$

This concludes the proof of Lemma 2. □

Corollary 1. *At the end of Phase 1 there are at least $n/8$ informed nodes, a.a.s.*

Proof. In Lemmas 1 and 2 we show that the number of informed nodes increases by a constant factor in every round. Hence, for suitable chosen α, $n/8$ nodes are informed after the $\alpha \log n$ steps of Phase 1. □

4.2 Phase 2

Recall that Phase 2 consists of $\alpha \log\log n$ many steps. In each step of the phase every informed node performs a push transmission. Lemma 3 shows that the number of uninformed nodes decreases by a constant factor in every round. Hence, at the end of Phase 2 there are at most $n/\log^5 n$ uninformed nodes. In this section we also present Lemma 4 which estimates, for any step of Phase 2, the number of nodes which are incident to at least one unused edge. This lemma will be used in the proof of our main results (see Sections 4.3.2 and 4.3.3).

Lemma 3 (Phase 2)**.** *Assume $\alpha \log n \leq t < \alpha(\log n + \log\log n)$ and $7n/8 \geq |H(t)| \geq n/(\log n)^{7\alpha}$. Then there exists a constant $c > 1$, independent of d and n, such that, w.h.p.,*

$$|H(t+1)| \leq |H(t)|/c.$$

Proof. Since $G_{n,d}$ is a random d-regular graph we can assume (see [1]) that for any $S \subset V$ with $|S| \leq n/2$ w.h.p. there exist two constants γ_1, γ_2 (depending on the size of S) with

$$\gamma_1 d \cdot |S| \;\leq\; |E(S, \overline{S})| \;\leq\; \gamma_2 d \cdot |S|.$$

Let $I_C(t) = \{v_1, v_2, \ldots, v_k\}$ and $H_C(t) = \{u_1, u_2, \ldots, u_\ell\}$, and let d_i denote the number of edges in $E(H(t), I(t))$ incident to u_i. Note that for $\alpha \log n \leq t \leq \alpha(\log n + \log\log n)$ every informed node performs only push transmissions (see Algorithm 1). Let us assume that every informed node only chooses one single neighbour in a step, and pushes the information to this neighbour only (as in the traditional push model). Let $X = |I^+(t)| = |H(t) \setminus H(t+1)|$ in the case when every informed node pushes the information to one single neighbour in a step. Then we have

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{j\in|H_C(t)|} 1-\left(1-\frac{1}{d}\right)^{d_j} \\
&= \sum_{j\in|H_C(t)|,d_j<d/2} 1-\left(1-\frac{1}{d}\right)^{d_j} + \sum_{j\in|H_C(t)|,d_j\geq d/2} 1-\left(1-\frac{1}{d}\right)^{d_j} \\
&\geq \sum_{j\in|H_C(t)|,d_j<d/2} \frac{d_j}{2d} + \sum_{j\in|H_C(t)|,d_j\geq d/2} 1-\frac{1}{\sqrt{e}} \\
&\geq \left(1-\frac{1}{\sqrt{e}}\right)\sum_{j\in|H_C(t)|} \frac{d_j}{d} \geq \frac{|E(H(t),I(t))|}{4d} \geq \gamma_1\cdot|H(t)|/4.
\end{aligned}$$

Here the first equality holds due to the fact that every edge in $E(H(t),I(t))$ is chosen for transmission with probability $1/d$ (we consider the traditional push model, not the push transmissions over four incident edges) and, hence, a node $u_i \in H_C(t)$ becomes informed with probability $1-(1-1/d)^{d_j}$. The first inequality holds due to the simple observation that for $1 \leq j \leq d/2$ we have

$$1-\left(1-\frac{1}{d}\right)^{j} \geq \frac{j}{2d}.$$

The second-to-last inequality holds since $\sum_{j\in|H_C(t)|} d_j = |E(H(t),I(t))|$.
Let G'_0 be the empty graph with vertex set $\{v_1,\ldots,v_k\}\cup H(t)$. For $1\leq i\leq k$ let G'_i denote the graph G'_{i-1} together with the edge used by v_i. Note that

$$k = |I_C(t)| \leq |H(t)|\cdot d = O(|H(t)|\cdot\log n).$$

Define $X_i = \mathbb{E}[X \mid G'_i]$. It is easy to verify that the sequence X_i is a Martingale, implying $X = X_0 = \mathbb{E}[X_k]$.
Now, since $|X_i - X_{i-1}| \leq 1$, this Martingale satisfies the bounded difference condition with value 1. Thus, we can apply the Azuma-Hoeffding inequality (Lemma 11) and obtain, for α large enough,

$$\Pr[|X_k-\mathbb{E}[X_k]| > \alpha\sqrt{\mathbb{E}[X_k]}\cdot\log n] \leq 2e^{-\left(\frac{\alpha^2\mathbb{E}[X_k]\log^2 n}{2k}\right)} = o(n^{-2}).$$

It remains to show that for $t=\alpha(\log n+\log\log n)$ the number of uninformed nodes is at least $n/(\log n)^{7\alpha}$. Note that in Phases 1 and 2, each node contacts four neighbours in at most $\alpha\log\log n+1$ steps (at most once in Phase 1 and at most $\alpha\log\log n$ times in Phase 2). Thus, we consider the following procedure. For $\alpha\log\log n+1$ steps *every* node chooses 4 distinct neighbours in each of these steps, and marks these nodes. Define Z as the number of unmarked nodes at the end of the above process. Note that for this process the number of marked nodes is larger than the number of informed nodes in the case of our protocol since we allow every node to mark neighbours. Hence $Z \leq |H(t+1)|$.
A node remains unmarked after one single step with probability $(1-4/d)^d$. Thus, a node remains unmarked after $\alpha\log\log n+1$ steps with a probability of

$$(1-4/d)^{d\alpha(\log\log n+1)} \geq (\log n)^{-6.9\alpha}.$$

This implies that $E[Z] \geq n/(\log n)^{-6.9\alpha}$.

We assume that the nodes for $G_{n,d}$ are numbered from 1 to n. Let G_0 be our original network with the unmarked nodes. Let G_i be the graph where all nodes are marked which are also marked in G_{i-1}, plus the nodes marked by the ith node with the process above. Define $Z_i = \mathbb{E}[Z \mid G_i]$. Again, is easy to verify that the sequence Z_i is a Martingale and $Z = Z_0 = E[Z_n]$. This Martingale satisfies the bounded difference condition with value 4. We can apply the Azuma-Hoeffding inequality from Lemma 11 and obtain

$$\Pr[|Z_k - \mathbb{E}[Z_n]| > n/(\log n)^{8\alpha}] \quad = \quad o(n^{-2}).$$

□

Corollary 2. *Assume $\alpha > 5\log_c 2$. At the end of Phase 2 all but at most $n/(\log n)^5$ nodes are informed, w.h.p.*

The next lemma estimates the size of the set $U(t) \subseteq V$, which is defined as the set of nodes incident to at least one edge which is not used before step $t+1$. This lemma will be used in the proofs of both of our main theorems.

Lemma 4. *(Phase 2) Let $\alpha \log n \leq t \leq \alpha(\log n + \log\log n)$. Then it holds w.h.p. that*

$$|U(t)| = \Omega\left(n\left(1 - \frac{1}{d}\right)^{10(t-\alpha\log n+1)}\right).$$

Proof. Recall that any node v having received the message during phase 1 will push the message exactly once during the $\alpha \log n$ steps of phase 1 (namely in the step immediately after it has received it), and it will push it during each of the $\alpha \log\log n$ steps of phase 2. Recall further than whenever a node pushes its message, it will select four nodes i.u.r. without replacement.
Consider values of t as specified in the statement of the lemma. We call an edge *t-busy* if it has been used at least once until step t; we call it *t-lazy* otherwise. We call a node *t-helpful* if it has at least one t-lazy edge (the rationale for naming it in such an apparently counter-intuitive way will become clear later).
We will estimate the number of t-helpful nodes. Unfortunately we cannot apply standard Chernoff bounds (like the one from Lemma 9 or Lemma 10) because the fact that a node is t-helpful is not independent of the helpfulness of other nodes (if we know that a node is t-helpful, i.e., it has at least one t-lazy edge, this implies a larger probability for its neighbours to be t-helpful as well).
Rather than arguing directly about nodes we shall consider edges first. In a single step, the probability for an arbitrary edge to be selected by one given neighbour during a single push operation is $4/d$. Thus the probability for it to be t-lazy, i.e., for never being used until step t, is at least $(1-4/d)^{2(t-\alpha\log n+1)}$. Unfortunately we cannot simply add these probabilities: there are dependencies between edges connected to the same node as well. We shall therefore focus on each node's *first* edge (first w.r.t. any arbitrary but fixed ordering). If the first edge of a node is t-lazy, then the node itself is helpful, and we obtain

$$\begin{aligned}
\mathbb{E}[|U(t)|] &\geq n \cdot \left(1 - \frac{4}{d}\right)^{2(t-\alpha\log n+1)} \\
&= n \cdot \left(1 - \frac{1}{d}\right)^{2(t-\alpha\log n+1)\cdot\frac{\log(1-4/d)}{\log(1-1/d)}} \\
&\geq n \cdot \left(1 - \frac{1}{d}\right)^{10(t-\alpha\log n+1)}.
\end{aligned}$$

(Notice that $4 \leq (\log(1-4/d))/(\log(1-1/d)) \leq 5$ for d large enough.) In order to conclude the proof, we again apply simple Martingale techniques. Let $V = \{v_1, \ldots, v_n\}$ and let $G_i'' = (V, E_i'')$ with $E_i'' = \{(v_j, v_\ell) \in E \mid 1 \leq j \leq i \text{ or } 1 \leq \ell \leq i\}$. That is, G_i'' consists of the graph of all nodes of V and all edges incident to the first i nodes of G. Then, $X_i = \mathbb{E}[|U(t)| \mid G_i'']$ is a Martingale. Since each node has d edges, the sequence $(X_i)_{i=0}^n$ satisfies the $(d+1)$-Lipschitz condition. Since $\mathbb{E}[|U(t)|] = n/\log^{O(1)} n$ for any $t < \alpha(\log n + \log\log n) + 1$ and $d \leq \delta \log n$, the Azuma-Hoeffding inequality from Lemma 11 yields the lemma. □

4.3 Remaining Phases

In this section we prove our main results for graphs with large degrees ($\delta \log\log n \leq d \leq \delta \log n$) or with small degrees ($\delta \leq d \leq \delta \log\log n$). To show the main results we will first apply the results of the previous section which estimate the number of uninformed nodes at the end of Phase 2 and the number of unused edges connected to these nodes. Then we will analyse the remaining phases of the algorithm. Recall that the remaining part of the algorithm (Phase 3 and Phase 4) is different for graphs with large or small degrees.
In the next section we first show some structural properties that we need for the analysis in Sections 4.3.2 and 4.3.3.

4.3.1 Structural Properties

To analyse the number of steps required to inform all nodes of the graph $G_{n,d}$ we use the following lemma whose proof is similar to the proof of Lemma 2.5 in [13]. The main purpose of the lemmas is to show that we can assume that the graph induced by the uninformed nodes can still be regarded as a random graph.
Before we state the lemma we need some additional definitions. We define $\mathcal{G}(d_1, \ldots, d_\ell)$ as the probability space of all graphs with ℓ nodes and degree sequence (d_i) with $d_i \leq d$. Hence, the ith node of any $G \in \mathcal{G}(d_1, \ldots, d_\ell)$ has degree d_i, $i \in \{1, \ldots, \ell - 1\}$. Recall that $G_{n,d} = (V, E)$ is random, i.e., the entries in its adjacency matrix are random variables taking values 0 or 1, according to the definition of d-regular random graphs with n vertices. For a fixed time step t, let $G(t) = (H(t), E(t))$ be the graph defined over the nodes $H(t)$ that are not informed in step t. We define $E(t) = \{(v, w) \mid (v, w) \in E \text{ and } v, w \in H(t)\}$. Note that $G(t)$ can be regarded as a random variable. For a *fixed* $S \subset V$ we define G_S as the *random* sub-graph of $G_{n,d}$ induced by the nodes of S, i.e. $G_S = (S, E \cap (S \times S))$. For fixed S we define $G_{\bar{S}} = (V, E_{\bar{S}})$ as a *fixed* sub-graph of $G_{n,d}$ that does not have any edges connecting two nodes in S, i.e., $E_{\bar{S}} = E \setminus (S \times S)$. Note that G_S is still a random graph, while $G_{\bar{S}}$ is fixed. Finally, we define $G_{n,d} \setminus G(t)$ as the graph $(V, E \setminus E(t))$. Note that $G_{n,d} \setminus G(t) = G_{\bar{H}(t)}$, the graph induced by the informed nodes.

For fixed t, $S \subset V$ and $G_{\bar{S}}$, let $A(t, S, G_{\bar{S}})$ be the event defined as $(H(t) = S) \wedge (G_{n,d} \setminus G_S = G_{\bar{S}})$. $A(t, S, G_{\bar{S}})$ is the event that the set S equals the set of uninformed nodes and that the original graph without the subgraph induced by the nodes of S equals $G_{\bar{S}}$. Note that $G_{\bar{S}}$ is a fixed graph over $V \setminus S$. If $H(t) = S$ then $G_S = G(t)$ and $G_{n,d} \setminus G_S = G_{\bar{H}(t)}$. Hence, $A(t, S, G_{\bar{S}})$ indicates if the set S equals the set of uninformed nodes and the graph induced by the informed nodes equals $G_{\bar{s}}$. $G(d_1, d_2, \ldots, d_{h(t)})$ is a random graph with degree sequence $d_1, d_2, \ldots, d_{h(t)}$, which is the degree sequence of the graph induced by the uninformed nodes. The following lemma now shows that we can assume that the graph $G(t)$ induced by the uninformed nodes is a random graph with respect to the degrees of these nodes.

Lemma 5. *If* $\Pr[A(t, S, G_{\bar{s}})] \neq 0$*, then for any fixed* $G(d_1, d_2, \ldots, d_{h(t)}) \in \mathcal{G}(d_1, \ldots, d_{h(t)})$*,*

$$\Pr\left[G(t) = G(d_1, d_2, \ldots, d_{h(t)}) \mid A(t, S, G_{\bar{S}})\right] = \frac{1}{|\mathcal{G}(d_1, \ldots, d_{h(t)})|}.$$

Proof. We assume w.l.o.g. that the vertices $v_1, \ldots, v_n \in V$ are ordered so that $v_1, \ldots, v_{h(t)} \in S$. For $1 \leq i, r \leq n$ we call an edge $(v_i, v_r) \in E$ the jth edge of v_i if there are exactly $j-1$ edges $(v_i, v_k) \in E$ with $k < r$. For $1 \leq i \leq n$ we define the event

$$B(v_i, j, \ell) = \{\text{node } v_i \text{ chooses its } j\text{th neighbour in step } \ell \leq t\}.$$

We define $U \subset V \times \{1, \ldots, n\} \times \{1, \ldots, t\}$ such that $|U \cap \{(v_i, j, \ell) \mid 1 \leq j \leq d|\}| \in \{0, 4\}$ for any $v_i \in V$ and $\ell \leq t$ (every node opens a channel to 4 or zero neighbours). Now we define

$$B(t) = \bigwedge_{(v_i, j, \ell) \in U} B(v_i, j, \ell).$$

Our goal is to show that for all G' the following probability is the same:

$$\Pr[G(t) = G' \mid (H(t) = S) \wedge (G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge B(t)].$$

This holds as long as

$$\Pr[G_{n,d} \setminus G_S = G_{\bar{S}} \wedge B(t)] \neq 0.$$

We know that

$$\begin{aligned} P_t &= \Pr[G(t) = G' \mid (H(t) = S) \wedge (G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge B(t)] \\ &= \frac{\Pr[(G(t) = G') \wedge (H(t) = S) \wedge (G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge B(t)]}{\Pr[(H(t) = S) \wedge (G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge B(t)]}. \end{aligned}$$

Assume that G' and G'' are two arbitrary graphs with $G', G'' \in \mathcal{G}(d_1, \ldots, d_{h(t)})$.
Now we show that if

$$(G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge (G_S = G') \wedge B(t)$$

results in $H(t) = S$, then the same holds for G''.
We prove by induction on i that $H(i)$ is the same in both $G' \cup G_{\bar{S}}$ and $G'' \cup G_{\bar{S}}$ for any $i \leq t$.
For $i = 0$ the assumption is trivially fulfilled. Now assume that the claim holds for $i-1$. If now a node v in the graph $G_{n,d} = G' \cup G_{\bar{S}}$ becomes informed in step i, then there must be some event $B(u, j, i)$ or $B(v, j', i)$ such that the jth edge of $u \in I(i-1)$ is adjacent to v, or the (j')th edge of v is adjacent to a node $u' \in I(i-1)$. In both cases the corresponding event implies that v in $G'' \cup G_{\bar{S}}$ becomes informed as well, since both edges (jth edge of u and (j')th edge of v) are contained in $G_{\bar{S}}$.
On the other hand, if a node v of $G' \cup G_{\bar{S}}$ is in $H(i)$, then for all events $B(u, j, i)$, for which the jth edge of u is adjacent to v, it holds that $u \in H(i-1)$. Similarly, for $B(v, j', i)$ we conclude that the (j')th edge of v is adjacent to some node $u' \in H(i-1)$. If now $u' \in V \setminus S$, then the (j')th edge of v is still adjacent to u' in $G'' \cup G_{\bar{S}}$, and v cannot be informed.
If the (j')th edge of v is adjacent to some $u' \in S$ in $G' \cup G_{\bar{S}}$, then v has fewer than $d - j'$ neighbours in $V \setminus S$, and the (j')th edge of v in $G'' \cup G_{\bar{S}}$ will be in G''. This finishes our induction.

Now assume

$$(G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge (G_S = G') \wedge B(t)$$

leads to $H(t) = S$. Then,

$$\begin{aligned} P_t &= \frac{\Pr[(G(t) = G') \wedge (H(t) = S) \wedge (G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge B(t)]}{\Pr[(H(t) = S) \wedge (G_{n,d} \setminus G_S = G_{\bar{S}}) \wedge B(t)]} \\ &= \frac{\Pr[(G_S = G') \wedge (G_{n,d} \setminus G_S = G_{\bar{S}})] \cdot \Pr[B(t)]}{\Pr[(G_{n,d} \setminus G_S = G_{\bar{S}})] \cdot \Pr[B(t)]} \\ &= \frac{\Pr[(G_S = G') \wedge (G_{n,d} \setminus G_S = G_{\bar{S}})]}{\Pr[(G_{n,d} \setminus G_S = G_{\bar{S}})]} \\ &= \Pr[(G_S = G') \mid (G_{n,d} \setminus G_S = G_{\bar{S}})] \\ &= \frac{1}{|\mathcal{G}(d_1, \ldots, d_{h(t)})|}. \end{aligned}$$

□

Since the set of uninformed nodes $H(t)$ induces $G(t)$, Lemma 5 implies that $G(t)$ equals *any* graph with its ($G(t)$'s) degree distribution with the same probability. The proof of the next lemma is obtained from Lemma 5. Recall that $A(t, S, G_{\bar{s}})$ indicates that the set S equals the set of uninformed nodes and the graph induced by the informed nodes equals $G_{\bar{S}}$. $G(d_1, d_2, \ldots, d_{h(t)})$ is a random graph with the same degree sequence as that of the graph induced by the uninformed nodes. The next lemma shows that, once the degree distribution of the uninformed nodes in $H(t)$ is given, the concrete position of the edges between the nodes in $H(t)$ is independent of the set of nodes $H(t+1)$ (which is the set of nodes that remain uninformed after one additional step).

Lemma 6. *Fix any $S' \subset S$. Then,*

$$\Pr\left[G(t) = G(d_1, d_2, \ldots, d_{h(t)}) \mid A(t, S, G_{\bar{S}}) \wedge H(t+1) = S'\right] = \frac{1}{|\mathcal{G}(d_1, \ldots, d_{h(t)})|},$$

as long as

$$\Pr\left[G(t) = G(d_1, d_2, \ldots, d_{h(t)}) \wedge A(t, S, G_{\bar{S}}) \wedge H(t+1) = S'\right] \neq 0.$$

Proof. Recall that in step $t+1$ all nodes perform a pull operation. We define the event $\mathcal{C}(t+1, S')$ as follows.

$$\begin{aligned} \mathcal{C}(t+1, S') &= (\text{In step } t+1, \text{ each } v' \in S' \text{ does not choose any stub connecting to } e \in G_{\bar{S}}, \text{ and} \\ &\quad \text{all nodes } v \in S \setminus S' \text{ choose at least one stub connecting } e \in G_{\bar{S}}) \wedge A(t, S, G_{\bar{S}}) \end{aligned}$$

Thus,

$$\begin{aligned} &\Pr\left[G(t) = G(d_1, d_2, \ldots, d_{h(t)}) \mid A(t, S, G_{\bar{S}}) \wedge H(t+1) = S'\right] \\ = \ &\Pr\left[G(t) = G(d_1, d_2, \ldots, d_{h(t)}) \mid A(t, S, G_{\bar{S}}) \wedge \mathcal{C}(t+1, S')\right]. \end{aligned} \tag{3}$$

Since the choices of the nodes in step $t+1$ are independent of the events $G(t) = G(d_1, d_2, \ldots, d_{h(t)})$ and $A(t, S, G_{\bar{S}})$, we can remove $\mathcal{C}(t+1, S')$ from (3), and the lemma follows. □

Let $H(t+1) \cap N(v) \neq \emptyset$ denote the event that v has at least one neighbour in $H(t+1)$. Furthermore, let $h_1(t)$ denote the number of nodes in $H(t)$ adjacent to at least one other node in $H(t)$. The lemma states that once $h_1(t)$ and $h(t)$ are given, a node in $H(t+1)$ has at least one neighbour in $H(t+1)$ with the same probability as in a random graph of size $h(t)$, where the random graph has the same degree distribution as $H(t)$.

Lemma 7. *For any node $v \in V$ the conditional probability*

$$\Pr\left[H(t+1) \cap N(v) \neq \emptyset \mid v \in H(t+1) \wedge h_1(t) = x \wedge h(t+1) = y\right]$$

is in the range

$$\left[\frac{y}{d \cdot x}, \frac{d^2 \cdot y}{x}\right].$$

Proof. Let $\mathcal{C}(d_1, \ldots, d_{|S|})$ denote the event that in some arbitrary but fixed set S the degree distribution of the nodes is given by $(d_1, \ldots, d_{|S|})$. Furthermore, we call two events A and B compatible if $\Pr[A \wedge B] \neq 0$. We define the following events.

$$\mathcal{B}(S, S', S''') = (H_1(t) = S \wedge H(t+1) = S' \wedge H(t) = S'')$$
$$\mathcal{H}(x, y) = (h_1(t) = x \wedge h(t+1) = y)$$

We define

$$\mathcal{A}(S, S', S'', G_{\bar{S}}, (d_1, \ldots, d_{|S''|}), G(d_1, \ldots, d_{|S''|})$$

as the union of the following events:

- $\mathcal{B}(S, S', S''')$,
- $\mathcal{C}(d_1, \ldots, d_{|S''|})$,
- $G(t) = G(d_1, \ldots, d_{|S''|})$, and
- $G_{n,d} \setminus G_S = G_{\bar{S}}$.

Then we have

$$
\begin{aligned}
&\Pr\left[H(t+1)\cap N(v)\neq\emptyset \mid v\in H(t+1)\wedge\mathcal{H}(x,y)\right]\\
&= \sum_{|S''|\geq x}\sum_{|S|=x}\sum_{|S'|=y}\sum_{G_{\bar{S}}}\sum_{(d_1,\ldots,d_{|S''|})}\sum_{G(d_1,\ldots,d_{|S''|})}\\
&\qquad \Pr\left[H(t+1)\cap N(v)\neq\emptyset \mid v\in H(t+1)\wedge\mathcal{A}(S,S',S'',G_{\bar{S}},(d_1,\ldots,d_{|S''|}),G(d_1,\ldots,d_{|S''|})\right]\cdot\\
&\qquad \Pr[v\in H(t+1)\wedge\mathcal{A}(S,S'S'',G_{\bar{S}}(d_1,\ldots,d_{|S''|}),G(d_1,\ldots,d_{|S''|})\mid\mathcal{H}(x,y)]\\
&= \sum_{|S''|\geq x}\sum_{|S|=x}\sum_{|S'|=y}\sum_{G_{\bar{S}}}\sum_{(d_1,\ldots,d_{|S''|})}\sum_{G(d_1,\ldots,d_{|S''|})}\\
&\qquad \Pr\left[H(t+1)\cap N(v)\neq\emptyset \mid v\in H(t+1)\wedge\mathcal{A}(S,S',S'',G_{\bar{S}},(d_1,\ldots,d_{|S''|}),G(d_1,\ldots,d_{|S''|})\right]\cdot\\
&\qquad \Pr\left[G(t)=G(d_1,\ldots,d_{|S''|})\mid\mathcal{B}(S,S',S''')\wedge\mathcal{C}(d_1,\ldots,d_{|S''|})\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\right]\cdot\\
&\qquad \Pr[\mathcal{B}(S,S',S''')\wedge C(d_1,\ldots,d_{|S''|})\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\mid\mathcal{H}(x,y)]\\
&= \sum_{|S''|\geq x}\sum_{|S|=x}\sum_{|S'|=y}\sum_{G_{\bar{S}}}\sum_{(d_1,\ldots,d_{|S''|})}\sum_{G(d_1,\ldots,d_{|S''|})}\\
&\qquad \Pr\left[H(t+1)\cap N(v)\neq\emptyset \mid v\in H(t+1)\wedge\mathcal{A}(S,S',S'',G_{\bar{S}},(d_1,\ldots,d_{|S''|}),G(d_1,\ldots,d_{|tS''|})\right]\cdot\\
&\qquad \Pr\left[G(t)=G(d_1,\ldots,d_{|S''|})\mid H(t+1)=S'\wedge H(t)=S''\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\right]\cdot\\
&\qquad \Pr[\mathcal{B}(S,S',S''')\wedge\mathcal{C}(d_1,\ldots,d_{|S''|})\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\mid\mathcal{H}(x,y)]\\
&\geq \frac{y}{d\cdot x}\cdot\sum_{|S''|\geq x}\sum_{|S|=x}\sum_{|S'|=y}\sum_{G_{\bar{S}}}\sum_{(d_1,\ldots,d_{|S''|})}\\
&\qquad \Pr[\mathcal{B}(S,S',S''')\wedge\mathcal{C}(d_1,\ldots,d_{|S''|})\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\mid\mathcal{H}(x,y)]=\frac{y}{d\cdot x}.
\end{aligned}
$$

In all the above sums we only sum up over compatible events. On the RHS of the first equality one can remove the term $v\in H(t+1)$ since the events are compatible, and

$$\mathcal{A}(S,S'S'',G_{\bar{S}}(d_1,\ldots,d_{|S''|}),G(d_1,\ldots,d_{|S''|})\Longrightarrow v\in H(t+1).$$

The third equality (before the first inequality) holds since the event $H(t)=S''$ together with $G_{n,d}\setminus G_S=G_{\bar{S}}$ completely defines the set of nodes in $H_1(t)$ as well as the degree distribution of the nodes in $H(t)$ (i.e., the event $C(d_1,\ldots,d_{|S''|})$).
Concerning the term

$$\Pr\left[G(t)=G(d_1,\ldots,d_{|S''|})\mid H(t+1)=S'\wedge H(t)=S''\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\right],$$

we know that all nodes in $S''\setminus S$ do not have any neighbours in $H(t)$ and thus

$$
\begin{aligned}
&\Pr\left[G(t)=G(d_1,\ldots,d_{|S''|})\mid H(t+1)=S'\wedge H(t)=S''\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\right]\\
=\;&\Pr\left[G(t)=G(d_1,\ldots,d_{|S''|})\mid H(t+1)=S'\wedge H(t)=S''\wedge G_{n,d}\setminus G_{S''}=G_{\bar{S''}}\right],
\end{aligned}
$$

where $G_{\bar{S''}}=G_{\bar{S}}$. According to the definition, $A(t,S'',G_{\bar{S''}})=(H(t)=S''\wedge G_{n,d}\setminus G_{S''}=G_{\bar{S''}})$, and hence Lemma 6 implies

$$\Pr\left[G(t)=G(d_1,\ldots,d_{|S''|})\mid H(t+1)=S'\wedge H(t)=S''\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\right]=\frac{1}{|\mathcal{G}(d_1,\ldots,d_{h(t)})|}.$$

Then,

$$
\begin{aligned}
&\sum_{G(d_1,\ldots,d_{|S''|})}\Pr\left[H(t+1)\cap N(v)\neq\emptyset \mid v\in H(t+1)\wedge\mathcal{A}(S,S',S'',G_{\bar{S}},(d_1,\ldots,d_{|S''|}),G(d_1,\ldots,d_{|S''|})\right]\cdot\\
&\qquad \Pr\left[G(t)=G(d_1,\ldots,d_{|S''|})\mid H(t+1)=S'\wedge H(t)=S''\wedge G_{n,d}\setminus G_S=G_{\bar{S}}\right]
\end{aligned}
$$

is the probability that a node in $S' = H(t+1)$ has a neighbour in S' in a random graph with degree distribution $(d_1, \ldots, d_{|S''|})$ defined over the vertices of S''. Since each node in $H_1(t)$ has at most d and at least 1 neighbour in $H(t)$, the probability above is at least $y/(d \cdot x)$.
The last equality follows from the fact that all the sums add up to 1. The upper bound is shown by replacing $y/(d \cdot x)$ with $d^2 \cdot y/x$ in the inequality. □

4.3.2 Phase 3 and Phase 4 for Graphs with Small Degrees

In this section we prove our main result for graphs with small degree (see Theorem 2).
In Phase 3 every informed node performs a pull transmission. Note that after that phase all nodes with fewer than four uninformed neighbours will themselves be informed. In Phase 4 all nodes that receive the message for the first time perform a push transmission. We will show that every node v that has (at the beginning of Phase 3) more than four uninformed neighbours is connected via a path of length $O(\log n/\log\log n)$ to a node itself informed during Phase 3. The path only contains nodes which have more than 4 uninformed neighbours at the beginning of Phase 4. This path can be used by the algorithm to inform node v in Phase 4.

Theorem 2. *Let δ be an arbitrary constant and let $\delta \le d \le \delta \log\log n$. Algorithm 1 informs all nodes of $G_{n,d}$ within $O(\log n)$ steps, a.a.s. Moreover, the number of message transmissions is bounded by $O(n \log\log n)$, a.a.s.*

Proof. Corollary 1 shows that, for α large enough, Algorithm 1 informs at least $n/8$ nodes during the first $\alpha \log n$ steps, w.h.p. Corollary 2 shows that, after an additional number of $\alpha \log\log n$ steps, all but at most $n/(\log n)^5$ nodes are informed whenever $\alpha > 5 \log_c 2$, where c is the constant defined in Lemma 3. Moreover, combining Lemma 3 with Lemma 4, we obtain

$$|U(t)| = \Omega(n(1-1/d)^{10(\alpha \log\log n)}) = \left(\log^{3+\Theta(1)} n\right) \cdot h(t)$$

at time $t = \alpha(\log n + \log\log n)$, whenever d and α are large enough. (That is, $c(1-1/d)^{10} > 1$ where c is the constant defined in Lemma 3, and
$\alpha > 3 \log_{c(1-1/d)^{10}} 2$. Notice that by letting d be large enough, we have $c(1-1/d)^{10} > 1$ since c is a constant which is independent of d.) According to Algorithm 1, in step $\alpha(\log n + \log\log n) + 1$ any informed node v sends the message to all the nodes calling v in this step (pull transmissions, Phase 3).
In order to prove the theorem, we build the graph $G_{n,d}$ by the following procedure. In each step $t \le \alpha(\log n + \log\log n)$, any node which performs a push transmission chooses four of its stubs and pairs them (if not paired yet) with unpaired stubs as per the configuration model [30]. Additionally, due to the definition of our algorithm, the uninformed vertices connected to these stubs become informed. In step $\alpha(\log n + \log\log n) + 1$, we pair all remaining stubs with each other. Let $S(t)$ denote the set of unpaired stubs at time t, and let

$$s(t) = |S(t)| \ge |U(t)|.$$

The outline of the proof is as follows. For $t = \alpha(\log n + \log\log n)$ (end of Phase 2) we show (see Lemma 8) that $h_1(t) = \Theta(h(t)^2 d^2/s(t))$ and $h_4(t) = \Theta(h(t)(h(t) \cdot d^2/s(t))^4)$, with probability $1 - e^{-\omega(\log^3 n)}$. Then we use that lemma and show below that with probability $1 - o(n^{-3})$ each node of $H_4(t)$ can be reached by a node of $H_1(t) \setminus H_4(t)$ via a path of length $O(\log n/\log\log n)$, using nodes of $H_4(t)$ only. Note that $H(t+1) \subseteq H_4(t)$ since in step $t+1$, nodes participate in pull transmissions, and only nodes with at least four uninformed neighbours can remain uninformed. Hence, nodes in $H(t+1) \subseteq H_4(t)$ will be informed by the end of Phase 3. Then we apply techniques

from e.g. [17] to consider the flow of information along one of the paths connecting nodes of $H_4(t)$ to nodes in $H_1(t) \setminus H_4(t)$. We conclude that within an additional $O(\log n)$ steps all nodes of the graph become informed, w.h.p. (Phase 4).

First we show that every node of $H_1(t) \setminus H_4(t)$ can be reached from a node of $H_1(t)$ via a path of length $O(\log n/\log\log n)$ containing only nodes from $H_4(t)$.
We know that every node of $H_4(t)$ has at least four neighbours in $H(t)$. We know from Lemma 8(2) and Lemma 8(1) that with probability $1-2^{-\omega(\log^3 n)}$,

$$\begin{aligned} h_4(t) &= \Theta\left(h(t)\cdot\left(\frac{h(t)\cdot d^2}{s(t)}\right)^4\right) = \Theta\left(\sqrt{\left(\frac{h_1(t)\cdot s(t)}{d^2}\right)}\cdot\left(\frac{(h_1(t))^2\cdot(s(t))^2\cdot d^8}{d^4\cdot(s(t))^4}\right)\right) \\ &= \Theta\left(\frac{(h_1(t))^{2.5}\cdot d^3}{(s(t))^{1.5}}\right) = \Theta\left(\frac{h_1(t)}{(\log n)^{9+O(1)}}\right). \end{aligned}$$

If $v_i(1),\dots,v_i(4)$ denote four neighbours of node $v_i \in H_4(t)$, then according to Lemma 5 all of these neighbours themselves are in $H_4(t)$ with probability at most

$$\left(\frac{h_4(t)\cdot d}{h_1(t)}\right)^4 \le \frac{1}{(\log n)^{36}}.$$

Claim: Let $L = O(\log n/\log\log n)$, and let u be a node at some distance $j \le L$ from a node $v \in H_4(t)$ in the graph induced by the nodes of $H(t)$. Furthermore, let $(v = u_0, u_1, \dots, u_j = u)$ be a shortest path between u and v in $H(t)$. Then with a probability of $1-n^{-1-\Omega(1)}$ it holds that

1. there is some u_ℓ with $\ell \le j$ which has a neighbour in $H_1(t) \setminus H_4(t)$, or
2. each u_ℓ, $\ell \le j$ has a set of neighbours $N_{u_\ell} = \{u_{\ell+1}, w_\ell^1, w_\ell^2, \dots\}$ in $H_4(t)$. There is at least one node in N_{u_ℓ} (one of them is denoted by $u_{\ell+1}$) with the property that the distance in $H(t)$ between v and such a node is exactly $\ell+1$. Additionally, there are at most two nodes in N_{u_ℓ} with a distance (in $H(t)$) of $\ell-1$ from v.

Note that the claim above can also be shown by extending a result of [6], which states that the nodes being at distance $O(\log\log\log n)$ from a node in a sparse random graph form either a tree or a uni-cyclic component. We first prove the following slightly reformulated claim and then we show how the original claim follows from the reformulated one. The reformulated claim is very similar to the original one, the only difference is the last sentence of Case 2, which now reads as follows: "Additionally, each such node in N_{u_ℓ} (i.e., with the property described above) has at most 2 neighbours in $H_4(t)$ at distance ℓ from v."
This reformulated claim is proved using induction on the distance j between u and v, where $j \le L+1$. For $j=0$ we have to consider v itself. Then, v either has a neighbour in $H_1(t)\setminus H_4(t)$, or it is connected with at least 4 stubs to some (not necessarily distinct) nodes of $H_4(t)$. Such a neighbour of v is connected with two more stubs to v with probability at most $(O(\log\log n)/(n/\log^{O(1)} n))^2 = n^{-1-\Omega(1)}$, and the claim holds for $j=0$.
Now we consider the inductive step from $j-1$ to j. Let u be a node at distance j from v. Then, according to the induction hypothesis, either a node u_ℓ with $\ell \le j-1$ has a neighbour in $H_1(t)\setminus H_4(t)$, or all nodes u_ℓ, $\ell \le j$, have at most two neighbours in $H_4(t)$ at distance $\ell-1$ from v.

In the first case we are done. Let us now concentrate on the second case. Since $u \in H_4(t)$, u has at least 2 additional neighbours in $H(t)$. Moreover, since $d = O(\log\log n)$, there are at most $(\log\log n)^{O(\log n/\log\log n)}$ nodes at distance $O(\log n/\log\log n)$ from v (each having at most d stubs pointing to nodes in $H(t)$) and we know that $h_1(t) = n/\log^{O(1)} n$. Thus, with probability

$$\left(\frac{(\log\log n)^{O(\log n/\log\log n)}}{n/\log^{O(1)} n}\right)^2 = n^{-1-\Omega(1)},$$

u has all of its neighbours from $H(t)$ at distance at most j from v. If now $w \in H(t)$ denotes a neighbour of u, where w is at distance $j+1$ from v, then w has two additional neighbours at distance j in $H(t)$ (and thus in $H_4(t)$) with probability at most

$$\left(\frac{(\log\log n)^{O(\log n/\log\log n)}}{n/\log^{O(1)} n}\right)^2 = n^{-1-\Omega(1)}.$$

If now $w \in H_1(t) \setminus H_4(t)$, then the first case of the reformulated claim holds, otherwise the second one holds. Hence, the reformulated claim follows.

It remains to show how the original claim follows from the reformulated claim. Let the sequence $(v = u_0, u_1, \ldots, u_j = u)$ denote a shortest path from u to v in the graph induced by the nodes of $H(t)$. According to the reformulated claim, there either is some u_ℓ, $\ell \leq j$, with a neighbour in $H_1(t) \setminus H_4(t)$, or all u_ℓ with $1 \leq \ell \leq j$ have at most two neighbours in $H_4(t)$ at distance $\ell - 1$ (this statement also holds for some neighbour w_j of u_j, which is at distance $j+1$ from v; however, this is not needed for the original claim). Moreover, each u_ℓ with $0 \leq \ell \leq j$ has at least one neighbour in $H_4(t)$ at distance $\ell + 1$ from v. This finishes the proof of the original claim.

Consider a path of length $\beta \log n/\log\log n$ with one endpoint in v. Then, each of the nodes on the path have all their neighbours in $H_4(t)$ with a probability of at most $\log^{-36\cdot\beta\log n/\log\log n} n = o(n^{-3})$, whenever β is large enough.

It remains to show that the information traverses the path within $\alpha \log n$ steps for a constant α large enough with $\alpha > \beta$. Let v be a node of $H_4(t)$, and let $u \in H_1(t) \setminus H_4(t)$ such that v and u are connected by a path $P_{u,v} = (u, v_1, \ldots, v_{k-1}, v_k = v)$ of length $k = \alpha \log n/\log\log n$ with $v_i \in H_4(t)$ for all $i \leq k-1$. Given that node v_i has the information at some time t, v_i transmits the information to v_{i+1} with probability $1/d \geq 1/O(\log\log n)$. Let now X_t denote the Bernoulli random variable which is 1 if and only if vertex v_i with

$$i = \max\{j \mid v_i \text{ is informed and } v_{j+1} \text{ does not have the information at time } t\}$$

transmits the information to v_{i+1} in step t. If v is already informed in step t then $X_t = 1$. We define $Y_1, \ldots, Y_{\alpha\log n}$ as Bernoulli random variables with $\Pr[Y_i = 1] = 1/d$. We define $Y = \sum_{t=1}^{\alpha\log n} Y_t$. Then $\mathbb{E}[Y] = \alpha\log n/d$. Clearly, the Y_t are independent and we may apply a Chernoff bound (Lemma 10):

$$\begin{aligned}\Pr[v \text{ is not informed}] &\leq \Pr\left[\sum_{t=1}^{\alpha\log n} X_t \leq k\right] \leq \Pr\left[\sum_{t=1}^{\alpha\log n} Y_t \leq k\right] \\ &\leq e^{-\frac{\alpha\log n}{2d}\cdot(1-\frac{\beta d}{\alpha\log\log n})} \leq o(n^{-3/\log\log n}).\end{aligned}$$

The last inequality holds for $\alpha\log\log n > \beta d$. Thus, v receives the information within $\alpha\log n$ steps, a.a.s. Note that with a more careful analysis the result can also be shown w.h.p.

To prove Theorem 2 it remains to bound the total number of messages. In Phase 1 every newly informed node performs a push transmission. The length of the phase is $\lceil \alpha \log n \rceil$, resulting in at most $O(n)$ many messages. Phase 2 has a length of $\alpha \log\log n$ steps, Phase 3 of one step, resulting in at most $n \log\log n$ many messages. Phase 4 takes $\alpha \log n$ many steps. Since only newly informed nodes communicate in that phase, and at the end of Phase 3 we have at $O(n/\log^5 n)$ uninformed nodes, the total number of messages is $O(n)$, resulting in a total of $O(n \log\log n)$ many messages. This finishes the proof of Theorem 2 □

In the remainder of this section we bound the sizes of the sets H_i.

Lemma 8. *With probability* $1 - e^{-\omega(\log^3 n)}$*, for* $t = \alpha(\log n + \log\log n)$ *we have:*

1. $h_1(t) = \Theta\left(\frac{(h(t))^2 \cdot d^2}{s(t)}\right)$,
2. $h_4(t) = \Theta\left(h(t) \cdot \left(\frac{h(t) \cdot d^2}{s(t)}\right)^4\right)$, *and*
3. $h_5(t) = \Theta\left(h(t) \cdot \left(\frac{h(t) \cdot d^2}{s(t)}\right)^5\right)$.

Proof. **Size of $\mathbf{H_1(t)}$.** We first calculate the expected number of these nodes and then apply Martingale techniques to prove the claim. We know that with probability $(h(t) \cdot d - 1)/s(t)$ a stub of a node $v \in H(t)$ is paired with a stub of a node of $H(t)$. Hence, the probability that there is a stub of v paired with a stub of $H(t)$ is

$$\Theta\left(\binom{d}{1} \cdot \left(\frac{(h(t) \cdot d - 1)}{s(t)}\right)\right) \cdot \Theta\left(\frac{h(t) \cdot d^2}{s(t)}\right).$$

Thus, if X is the random variable denoting the number of nodes in $H(t)$ which have a neighbour in $H(t)$, then

$$\mathbb{E}[X] = \Theta\left(\frac{h(t)^2 \cdot d^2}{s(t)}\right) = \frac{n}{(\log n)^{O(1)}}.$$

If $H(t) = \{v_1, \ldots, v_{h(t)}\}$, let G'_i be the graph induced by the vertices $\{v_1, \ldots, v_i\} \cup \bigcup_{j=1}^{i} N(v_j)$, where $N(v_j)$ is the set of neighbours of v_j. Furthermore, let $X_i = \mathbb{E}[X \mid G'_i]$ be the corresponding Martingale sequence. It is easy to verify that X_i satisfies the $(d+1)$-Lipschitz condition. Applying the Azuma-Hoeffding inequality (Lemma 11) we obtain

$$\Pr[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]/2] \leq e^{-\Theta((h(t)^2 d^2/s(t))^2/(d^2 h(t)))} \leq 2^{-\omega(\log^3 n)}$$

since $h(t) \leq n/\log^{2+\Theta(1)} n$ and $d \leq \delta \log n$.

Size of $\mathbf{H_4(t)}$. This proof is similar to the one for $H_1(t)$. Assume again that $v \in H(t)$. With probability

$$\begin{aligned}
&\Theta\left(\binom{d}{4} \cdot \left(\frac{(h(t) \cdot d - 1}{s(t)}\right) \cdot \left(\frac{(h(t) \cdot d - 2}{s(t) - 1}\right) \cdot \left(\frac{(h(t) \cdot d - 3}{s(t) - 2}\right) \cdot \left(\frac{(h(t) \cdot d - 4}{s(t) - 3}\right)\right) \\
&= \Theta\left(\left(\frac{(h(t) \cdot d^2}{s(t)}\right)^4\right)
\end{aligned}$$

there are four stubs of v paired with stubs of nodes in $H(t)$. Thus, if Y is the random variable denoting the number of nodes in $H(t)$ which have at least four neighbours in $H(t)$, then

$$\mathbb{E}[Y] = \Theta\left(h(t) \cdot \left(\frac{h(t) \cdot d^2}{s(t)}\right)^4\right) = \frac{n}{(\log n)^{O(1)}}.$$

Similar to the analysis of $\mathbf{H_1(t)}$, define the Martingale sequence $Y_i = \mathbb{E}[Y \mid G_i']$. The Y_i satisfy the $(d+1)$-Lipschitz condition, and the claim follows from an application of Azuma-Hoeffding bounds (Lemma 11), as in the previous case.

Size of $\mathbf{H_5(t)}$: With probability

$$\Theta\left(\frac{(h(t) \cdot d^2}{s(t)}\right)^5,$$

there are five stubs of v paired with stubs of $H(t)$. Thus, if Y is the random variable denoting the number of nodes in $H(t)$ which have at least four neighbours in $H(t)$, then

$$\mathbb{E}[Y] = \Theta\left(h(t) \cdot \left(\frac{h(t) \cdot d^2}{s(t)}\right)^5\right) = \frac{n}{(\log n)^{O(1)}}.$$

Similar to the analysis of $\mathbf{H_1(t)}$, define the Martingale sequence $Y_i = \mathbb{E}[Y \mid G_i']$. The Y_i satisfy the $(d+1)$-Lipschitz condition, and the claim follows from an application of (Lemma 11), as in the two cases above. □

Observation 1. *Assume $s(t) = \Theta(nd)$ for the same t as in Lemma 8. Then, with probability at least $1 - e^{-\omega(\log^3 n)}$,*

1. $h_1(t) = \Theta\left(\frac{(h(t))^2 \cdot d}{n}\right)$,
2. $h_4(t) = \Theta\left(h(t) \cdot \left(\frac{h(t) \cdot d}{n}\right)^4\right)$, *and*
3. $h_5(t) = \Theta\left(h(t) \cdot \left(\frac{(h(t) \cdot d}{n}\right)^5\right)$.

Proof. For $s(t) = \Theta(nd)$, all three equations follow directly from Lemma 8. □

4.3.3 Phase 3 for Graphs with Large Degrees

Now we turn our attention to the case $d \geq \delta \log\log n$ (but $d \leq \delta \log n$) where δ is a large constant. In this case our algorithm has only one more phase. In Phase 3 every informed node performs a pull transmission. The length of the phase is $\alpha \log\log n$. To prove Theorem 3 (i.e., the result for large degrees), we first estimate the one-step decrease of the set of uninformed nodes in Claim 1. The correctness of the algorithm now follows from applying Claim 1 for every step of the phase.

Theorem 3. *Let δ be a sufficiently large constant and let $\delta \log\log n \leq d \leq \delta \log n$. Algorithm 2 broadcasts a message in $G_{n,d}$ a.a.s. within $O(\log n)$ steps, and the number of message transmissions is bounded by $O(n \log\log n)$.*

Proof. Again we can apply Corollary 1 showing that the algorithm informs at least $n/8$ nodes in Phase 1 (the first $\alpha \log n$ steps). At the end of Phase 2 all but $n/(\log n)^{\Theta(1)}$ nodes are informed (see Corollary 2). Recall that $U(t)$ is the set of nodes incident to at least one edge which is not used before step $t+1$. Lemma 4 implies that

$$|U(t)| = \Omega\left(n \cdot \left(1 - \frac{1}{d}\right)^{10\alpha \log\log n}\right) = (\log n)^{5+\Theta(1)} \cdot h(t)$$

at time $t = \alpha(\log n + \log\log n)$, whenever $c(1-1/d)^{10} > 1$ and $\alpha > 6 \log_{c(1-1/d)^{10}} 2$. Here c is the constant defined in Lemma 3.
According to the algorithm described above, in any step $t+i$, where $i > 0$, any informed node v transmits the message to all nodes which call v in this step. We show the following two claims in a single induction. The claims estimate the one-step decrease of the set of uninformed nodes.

Claim 1: As long as $h(t+i+2) \geq (\log n)^q$ with q being a large constant,

$$\frac{h(t+i+2)}{h(t+i+1)} = o\left(\frac{h(t+i+1)}{h(t+i)}\right)^2, \tag{4}$$

w.h.p.

Claim 2: For any $i \geq 1$

$$h_1(t+i) \leq \frac{h(t+i-1)}{d^3} \tag{5}$$

with probability $1 - e^{-\omega(\log^3 n)}$.

Induction base, Claim 1: In order to show that the first claim holds for $i = 0$ we assume that the information is only transmitted by pull transmissions. Recall that $S(t)$ denotes the set of unpaired stubs at time t, and $s(t) = |S(t)|$. We know that if $\delta \gg \alpha$, then $s(t) = \Theta(nd)$. Hence, we can use Observation 1 to show bounds on $h_1(t)$, $h_4(t)$, and $h_5(t)$.
All nodes $v \in H_3(t) \setminus H_4(t)$ will become informed due to the `pull(`$\mathcal{M}$`)` in step $t+1$. Only nodes with four or more uninformed neighbours can remain uninformed. A node $v \in H_4(t) \setminus H_5(t)$ remains uninformed due to the `pull(`$\mathcal{M}$`)` in step $t+1$ with probability $1/\binom{d}{4}$, independent of the other nodes. Hence, the expected number of nodes in $H_4(t) \setminus H_5(t)$ that remain uninformed is

$$\frac{h_4(t)}{\binom{d}{4}} = \Theta\left(h(t) \cdot \left(\frac{h(t) \cdot d^2}{s(t)}\right)^4\right) \cdot \frac{1}{d^4}.$$

We can apply Chernoff bounds from Lemma 10 to conclude that with probability $1 - e^{-\omega(\log^3 n)}$,

$$
\begin{aligned}
h(t+1) &\leq \left(h(t)\cdot\left(\frac{h(t)\cdot d^2}{s(t)}\right)^4\cdot\frac{1}{d^4}\right)\cdot(1\pm o(1))+h_5(t)\\
&\leq \left(h(t)\cdot\left(\frac{h(t)\cdot d}{s(t)}\right)^4\right)\cdot(1\pm o(1))+\Theta\left(h(t)\cdot\left(\frac{h(t)\cdot d^2}{s(t)}\right)^5\right) \qquad (6)\\
&\leq \left(h(t)\cdot\left(\frac{h(t)\cdot d}{s(t)}\right)^4\right)\cdot(1\pm o(1))+\Theta\left(h(t)\cdot\left(\frac{h(t)\cdot d}{s(t)}\right)^4\right)\cdot\left(\frac{h(t)\cdot d^6}{s(t)}\right) \qquad (7)\\
&\leq \left(h(t)\cdot\left(\frac{h(t)\cdot d}{s(t)}\right)^4\right)\cdot(1\pm o(1)). \qquad (8)
\end{aligned}
$$

The last inequality holds since $h(t)\leq n/(\log n)^7$, w.h.p., and $d\leq\log n$ (see Lemma 3).
Now we consider the set $H(t+2)$. We know from Lemma 5 that the sub-graph induced by the vertices of $H(t)$ is a random graph with degree sequence $(d_1,\ldots,d_{h(t)})$, where d_i denotes the number of neighbours of node $v_i\in H(t)$. Similar to the proof of Lemma 8 we can show that with probability $1-e^{-\omega(\log^3 n)}$,

$$
h_1(t+1)=\Theta\left(h(t+1)\cdot\frac{h(t+1)}{h_1(t)}\right) \qquad (9)
$$

and

$$
h_4(t+1)=\Theta\left(h(t+1)\cdot\left(\frac{h(t+1)}{h_1(t)}\right)^4\right). \qquad (10)
$$

From Observation 1 and Equation 8 it follows that with probability $1-e^{-\omega(\log^3 n)}$,

$$
\begin{aligned}
(h(t+1))^2 &= \left(\frac{h_4(t)}{d^4}\right)^2\cdot(1\pm o(1))=\left(\frac{\left(h(t)\cdot\left(\frac{h(t)\cdot d}{n}\right)^4\right)}{d^4}\right)^2\cdot(1\pm o(1))\\
&= \frac{(h(t))^{10}}{n^8}\cdot(1\pm o(1))=o\left(\frac{h(t)^6\cdot d^4}{n^4}\right). \qquad (11)
\end{aligned}
$$

The last inequality holds since $h(t)\leq n/(\log n)^{\Theta(1)}$, $d\leq\delta\log n$ and $(h(t))^4/n^4\ll d^4$.
Since $H(t+2)\subset H_4(t+1)$, we can use Equation 10 and then Observation 1, resulting in

$$\begin{aligned}
\frac{h(t+2)}{h(t+1)} &\leq \frac{h_4(t+1)}{h(t+1)} = O\left(\frac{h(t+1)\cdot\left(\frac{h(t+1)}{h_1(t)}\right)^4}{h(t+1)}\right) \\
&= O\left(\left(\frac{h(t+1)}{\frac{h(t)^2\cdot d}{n}}\right)^4\right) = O\left(\left(\frac{h(t+1)}{h(t)}\right)^4\cdot\left(\frac{1}{h(t)\cdot\frac{d}{n}}\right)^4\right) \\
&= O\left(\left(\frac{h(t+1)}{h(t)}\right)^4\cdot\left(\frac{(\log n)^{\Omega(1)}}{n\cdot\frac{d}{n}}\right)^4\right) \\
&= O\left(\left(\frac{h(t+1)}{h(t)}\right)^2\cdot\left(\frac{h(t+1)}{h(t)}\right)^2\cdot\left(\frac{(\log n)^{\Omega(1)}}{d}\right)^4\right) \\
&= o\left(\left(\frac{h(t+1)}{h(t)}\right)^2\cdot\left(\frac{h(t)\cdot d}{n}\right)^4\cdot\left(\frac{(\log n)^{\Omega(1)}}{d}\right)^4\right) \\
&= o\left(\frac{h(t+1)^2}{h(t)^2}\right).
\end{aligned}$$

This shows that Claim 1 holds for $i = 0$.

Induction base, Claim 2: From Lemma 8 we know that

$$h_1(t+i) = \Theta\left(\frac{(h(t+i))^2\cdot d^2}{s(t+i)}\right).$$

We also know that $h(t+i)/s(t+1) = O(d^5)$ and $h(t+1) \leq h(t+i-1)$. Hence,

$$h_1(t+i) \leq \frac{h(t+i)}{d^3} \leq \frac{h(t+i-1)}{d^3}. \tag{12}$$

Assume now that Claim 1 and Claim 2 hold for some $i - 1 \geq 0$.

Inductive step, Claim 1: Similar to Equation 9 we can use Lemma 7 to obtain

$$\mathbb{E}[h_1(t+i)] > h(t+i)\cdot\frac{h(t+i)}{d\cdot h_1(t+i-1)}.$$

Now we can use Claim 2 (Equation 12) and get

$$\mathbb{E}[h_1(t+i)] \geq \frac{(h(t+i))^2\cdot d^2}{h(t+i-1)}. \tag{13}$$

Applying simple Martingale techniques as before, we conclude that with probability $1 - e^{-\omega(\log^3)}$

$$h_1(t+i) > \mathbb{E}[h_1(t+i)](1-o(1)). \tag{14}$$

Similarly, a calculation like the one for Equation 10 together with Lemma 7 results in

$$\mathbb{E}[h_4(t+i+1)] \leq h(t+i+1) \cdot \left(\frac{h(t+i+1) \cdot d}{h_1(t+i)}\right)^4. \tag{15}$$

Using Equation 13 we get

$$\begin{aligned}
\mathbb{E}[h_4(t+i+1)] &\leq h(t+i+1) \left(\frac{h(t+i+1) \cdot d}{\frac{h(t+i)^2 \cdot d^2}{h(t+i-1)} \cdot (1-o(1))}\right)^4 \\
&\leq h(t+i+1) \left(\frac{\frac{h(t+i+1)}{h(t+i)}}{d \cdot \frac{h(t+i)}{h(t+i-1)}}\right)^4 \cdot (1+o(1)) \\
&\leq h(t+i+1) \left(\frac{h(t+i+1) \cdot h(t+i-1)}{d \cdot (h(t+i))^2}\right)^4 \cdot (1+o(1)).
\end{aligned}$$

The Induction hypothesis gives us

$$h(t+i+1) = O\left(\frac{h(t+i)^3}{h(t+i-1)^2}\right),$$

resulting in

$$\begin{aligned}
\mathbb{E}[h_4(t+i+1)] &\leq h(t+i+1) \left(\frac{h(t+i+1) \cdot h(t+i-1)}{d \cdot (h(t+i))^2}\right)^4 \cdot (1+o(1)) \\
&\leq h(t+i+1) \cdot o\left(\frac{h(t+i+1)^2}{h(t+i)^2}\right).
\end{aligned} \tag{16}$$

Applying similar Martingale techniques as before, we conclude that with probability $1 - e^{-\omega(\log^3 n)}$,

$$h_4(t+i+1) = \mathbb{E}[h_4(t+i+1)](1 \pm o(1)). \tag{17}$$

Since $h_4(t+i+1) \geq h(t+i+2)$ we get from Equations 16 and 17 that

$$h(t+i+2) \leq h(t+i+1) \cdot o\left(\frac{h(t+i+1)^2}{h(t+i)^2}\right),$$

which proves the inductive step for Claim 1.

Inductive step, Claim 2: We know that each node v of $H(t+i)$ has at least 4 (and at most d) edges to nodes in $H_1(t+i-1)$ (otherwise, v would have been informed in step $t+i$). Wach node in $H_1(t+i-1)$ has at least one neighbour in $H_1(t+i-1)$. Thus, applying Lemma 7 we get that v does not have any neighbour in $H(t+i)$ with probability at least

$$\left(1 - \frac{d \cdot h(t+i)}{h_1(t+i-1)}\right)^d \geq 1 - \frac{h(t+i) \cdot d^2}{h_1(t+i-1)}.$$

Then,

$$\mathbb{E}[h_1(t+i)] < \frac{h(t+i)^2 \cdot d^2}{h_1(t+i-1)}. \tag{18}$$

The probability that v has a neighbour in $H(t+i)$ is larger than

$$\frac{h(t+i)}{h_1(t+i-1) \cdot d}.$$

From Equation 5 we get $h_1(t+i-1) \leq \frac{h(t+i-2)}{d^3}$. Hence,

$$\begin{aligned}\mathbb{E}[h_1(t+i)] &> h(t+i) \cdot \frac{h(t+i)}{h_1(t+i-1) \cdot d} \\ &\geq h(t+i) \cdot \frac{h(t+i) \cdot d^2}{h(t+i-2)}.\end{aligned} \tag{19}$$

Now we consider the stubs of nodes in $H(t+1)$ one by one, and apply the following process. Assume these stubs are numbered $1, 2, \ldots, s$, where s is the number of stubs in $H(t+i)$. In some step i, we pair the ith *free* stub with one of the stubs not paired so far. Clearly, it follows from our induction hypothesis that

$$h_1(t+i-1) = \omega(h(t+i) \cdot d).$$

Then, in each step at most 2 stubs from $H(t+i)$ are paired with each other, and in step $i < s/4$, the ith free stub is paired with a stub in $H(t+i)$ with some probability in the range

$$\left[\frac{s}{2d \cdot h_1(t+i-1)}, \qquad \frac{s \cdot d^2}{h_1(t+i-1) - s}\right],$$

regardless of the connections established so far. Moreover, the upper bound holds even for $i \geq s/4$. Taking into account that (due to our assumption) $h(t+i) > \log^q n$, we apply Lemma 10 to conclude that the number of stubs matched by this process is between

$$\Omega\left(\frac{s^2}{(d \cdot h_1(t+i-1)}\right) \text{ and } O\left(\frac{s^2 \cdot d^2}{h_1(t+i-1)}\right),$$

with probability $1 - e^{-\omega(\log^3 n)}$. Since each node in $H(t+i)$ has more than one and at most d neighbours in $H_1(t+i-1)$, with probability $1 - e^{-\omega(\log^3 n)}$ the number of nodes in $H_1(t+i)$ is between

$$\Omega\left(\frac{h(t+i)^2}{d^2 \cdot h_1(t+i-1)}\right) \text{ and } O\left(\frac{h(t+i)^2 d^2}{h_1(t+i-1)}\right). \tag{20}$$

Now we can use Equation 20 (with a suitable chosen constant c) and Claim 1:

$$
\begin{aligned}
\frac{h_1(t+i)}{h(t+i)} &\leq \frac{h(t+i)^2 \cdot d^2}{h(t+i) \cdot h_1(t+i-1)} \leq \frac{h(t+i) \cdot d^2}{h_1(t+i-1)} \\
&\leq \frac{h(t+i) \cdot d^2}{c \cdot \frac{h(t+i-1)^2}{(d^2 \cdot h_1(t+i-2)}} && (21) \\
&\leq \frac{h(t+i) \cdot d^2}{\frac{c \cdot h(t+i-1)^2 \cdot d}{h(t+i-2)}} && (22) \\
&\leq \frac{h(t+i) \cdot d \cdot h(t+i-2)}{c \cdot (h(t+i-1))^2} && (23) \\
&\leq \frac{h(t+i-1) \cdot d}{c \cdot h(t+i-2)}. && (24)
\end{aligned}
$$

To obtain Eq. 21 we used the induction hypothesis and to obtain 24 we used Claim 1, respectively. Since

$$\frac{h(t+i-1) \cdot d}{h(t+i-2)} \leq \frac{h(t+1) \cdot d}{h(t)}$$

and

$$h(t+1) = o(h(t)/d^4)$$

(see above), we obtain the inductive step of Claim 2.

The correctness of the algorithm now follows from a $(\alpha \log\log n)$-fold application of Claim 1. If $h(t) \leq \log^q n$, then the last nodes become informed within $O(\log\log n)$ steps, a.a.s. [35]. It remains to bound the total number of messages. Similar to the proof of Theorem 2, Phases 1 and 2 use at most $(On \log\log n)$ many messages. Phase 3 has a length of $\alpha \log\log n)$ resulting in at most $O(n \log\log n)$ many messages. This finishes the proof of Theorem 3. □

5 Conclusions

We considered a simple modification of the random phone call model in d-regular random graphs where each node contacts four *distinct* neighbours in every time step. We showed that this modification leads to a significant improvement in the number of transmissions required for broadcasting a message to all nodes of the graph. One interesting question is how much randomness is needed in the graph to obtain the improvements described above. We know that on graphs with similar expansion and connectivity properties as in d-regular random graphs the models presented above may not lead to any notable improvement. An example for such a graph is the Cartesian product of a d-regular random graph with a K_5. Another important question is whether four choices are necessary. We believe that the same results may also be obtained with three choices. However, the case of two random choices per time step is still completely open.

6 Acknowledgements

The authors are deeply grateful to the anonymous reviewers for their patience and their most helpful comments and remarks.

References

[1] B. Bollobás. *Random Graphs*. Academic Press, 1985.

[2] S.M. Botros and S.R. Waterhause. Search in JXTA and other distributed networks. In *Proc. of P2P'01*, pp. 30–35, 2001.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, Randomized Gossip Algorithms. *In IEEE Transactions on Information Theory and IEEE/ACM Transactions on Networking*, 52:2508–2530, 2006.

[4] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *The Annals of Mathematical Statistics*, 23:493–507, 1952.

[5] C. Cooper, M. Dyer, and C. Greenhill. Sampling regular graphs and a peer-to-peer network. In *Proc. of SODA'05*, pp. 980–988, 2005.

[6] C. Cooper and A. Frieze The cover time of sparse random graphs. *Random Structures and Algorithms*, 30:1–16, 2007.

[7] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated database maintenance. In *Proc. of PODC'87*, pp. 1–12, 1987.

[8] B. Doerr, M. Fouz, T. Friedrich. Social networks spread rumors in sublogarithmic time In *Proc. of STOC'11*, pp. 21–30, 2011.

[9] B. Doerr, T. Friedrich, T. Sauerwald. Quasirandom Rumor Spreading In *Proc. of SODA'08*, pp. 773–781, 2008.

[10] D. Dubhashi and A. Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2009.

[11] R. Elsässer. On the communication complexity of randomized broadcasting in random-like graphs. In *Proc. of SPAA'06*, pp. 148–157, 2006.

[12] R. Elsässer and T. Sauerwald. Broadcasting vs. mixing and information dissemination on Cayley graphs. In *Proc. of STACS'07*, pp. 163–174, 2007.

[13] R. Elsässer and T. Sauerwald. The power of memory in randomized broadcasting. In *Proc. of SODA'08*, pp. 290–227, 2008.

[14] P. Erdős and A. Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[15] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[16] T. Feder, A. Guetz, M. Mihail, and A. Saberi. A local switch Markov chain on given degree graphs with application in connectivity of peer-to-peer networks. In *Proc. of FOCS'06*, pp. 69–76, 2006.

[17] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Random Structures and Algorithms*, 1(4):447–460, 1990.

[18] J. Friedman. A proof of Alon's second eigenvalue conjecture. In *Proc. of STOC'03*, pp. 720–724, 2003.

[19] A.M. Frieze and G.R. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10:57–77, 1985.

[20] N. Fountoulakis and K. Panagiotou. Rumor Spreading on Random Regular Graphs and Expanders. In Proceedings of the 14th International Workshop on Randomization and Computation (RANDOM '10), pp. 560-573, 2010.

[21] Gnutella. *The gnutella protocol specification v.0.4.*

[22] T. Hagerup and C. Rüb. A guided tour of Chernoff bounds. *Information Processing Letters*, 36(6):305–308, 1990.

[23] S. Hoory, N. Linial, and A. Wigderson. Expander Graphs and Their Applications. *Bulletin of the AMS*, 43(4):439–561, 2006.

[24] S. Jagannathan, G. Pandurangan, and S. Srinivasan. Query protocols for highly resilient peer-to-peer networks. In *Proc. of ISCA PDCS'06*, pp. 247–252, 2006.

[25] R. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized rumor spreading. In *Proc. of FOCS'00*, pp. 565–574, 2000.

[26] D. Kempe, J. Kleinberg, and A. Demers. Spatial gossip and resource location protocols. In Proceedings of the 33rd Symposium on the Theory of Computing (*STOC*), pp. 163–172, 2001.

[27] C. Law and K.-Y. Siu. Distributed construction of random expander networks. In Proceedings of the 22nd *INFOCOM*, pp. 2133–2143, 2003.

[28] R. Motwani and P. Raghavan *Randomized Algorithms* Cambridge University Press, 1995.

[29] P. Mahlmann and C. Schindelhauer. Distributed random digraph transformations for peer-to-peer networks. In *Proc. of SPAA'06*, pp. 308–317, 2006.

[30] B.D. McKay and N.C. Wormald. Asymptotic enumeration by degree sequence of graphs with degrees o(sqrt(n)). Combinatorica, 11:369–382, 1991.

[31] M. Newman. The Structure and Function of Complex Networks. SIAM Review, 45(2):167–256, 2003.

[32] G. Pandurangan, P. Raghavan, and E. Upfal. Building low-diameter peer-to-peer networks. In *Proc. of FOCS'01*, pp. 492–499, 2001.

[33] B. Pittel. On spreading a rumor. SIAM Journal on Applied Mathematics, 47(1):213–223, 1987.

[34] R.W. Robinson and N.C. Wormald. Almost all regular graphs are Hamiltonian. Random Structures and Algorithms, 5:363–374, 1994.

[35] T. Sauerwald. On Mixing and Edge Expansion Properties in Randomized Broadcasting . In *Proc. of ISAAC'07*, pp. 196–207, 2007.

A Tail Bounds

First we present two different versions of Chernoff bounds.

Lemma 9 ([22]). *Let $X = X_1 + X_2 + \ldots + X_n$ be independent $0-1$ random variables with $\Pr[X_i = 1] = p$ and $\Pr[X_i = 0] = 1 - p$. Then for $0 < \alpha < 1$ and $a \geq p$ we have*

$$\Pr[X \geq \alpha n] \leq \left(\frac{p}{\alpha}\right)^{\alpha} \cdot \left(\frac{1-p}{1-\alpha}\right)^{1-\alpha}.$$

Lemma 10 ([4]). *Let $X = X_1 + X_2 + \ldots + X_n$ be independent $0-1$ random variables with $\Pr[X_i = 1] = p$ and $\Pr[X_i = 0] = 1 - p$.*

1. *For $\epsilon > 0$ we have $\Pr[X \geq (1+\epsilon) \cdot E[X]] \leq \left(\frac{e^{\epsilon}}{(\epsilon+1)^{\epsilon+1}}\right)^{E[X]}$.*
2. *For $0 < \epsilon < 1$ we have $\Pr[X \geq (1+\epsilon) \cdot E[X]] \leq e^{-\frac{\epsilon^2 E[X]}{3}}$.*

The following well-known bound is called Azuma-Hoeffding inequality.

Lemma 11 ([10]). *Let $X_0, X_1, \ldots$ be a Martingale sequence satisfying the Lipschitz-condition so that $|X_i - X_{i-1}| \leq c_i$ for a constant c_i. Let $c = c_0 + c_1 + \ldots + c_n$ Then*

1. $\Pr[X_n \geq X_0 + t] \leq e^{\frac{-t^2}{2c}}$.
2. $\Pr[X_n \leq X_0 - t] \leq e^{\frac{-t^2}{2c}}$.